

# Differential Zernike Filter for Phasing of Segmented Mirror and Image Processing

Tao Huang

Submitted for the Degree of Doctor of Philosophy

Heriot-Watt University

Department of Physics

Feb. 2009

The copyright in this thesis is owned by the author. Any quotation from the thesis or use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information.

## **Abstract**

The major objective of this thesis is to study the differential Zernike filter and its applications in phasing segmented mirror and image processing. In terms of phasing, we provide both theoretical analysis and simulation for a differential Zernike filter based phasing technique, and find that the differential Zernike filter perform consistently better than its counterpart, traditional Zernike filter. We also combine the differential Zernike filter with a feedback loop, to represent a gradient-flow optimization dynamic system. This system is shown to be capable of separating (static) misalignment errors of segmented mirrors from (dynamical) atmospheric turbulence, and therefore compress the effects of atmospheric turbulence. Except for segmented mirror phasing, we also apply the Zernike feedback system in image processing. For the same system dynamics as well as in segment phasing, the Zernike filter feedback system is capable of compress the static noisy background, and makes the single particle tracking algorithm even working in case of very low signal-to-noise ratio. Finally, we apply an efficient multiple-particle tracking algorithm on a living cell image sequence. This algorithm is shown to be able to deal with higher particle density, while the single particle tracking methods are not working under this condition.

## **Acknowledgements**

It is my pleasure to thank my supervisor Dr Weiping Lu for his continual daily support and sharing of his knowledge, this thesis has benefited greatly from his most helpful comments and constructive criticism.

Thank Dr. Alan H. Greenaway for fruitful discussion and his great help for my published paper in Heriot-Watt University.

I gratefully acknowledge the former member of Waves & Fields Group of Heriot-Watt University, Dr. Sijiong Zhang, for his great encouragement and support during my time in Edinburgh even after he moved to Cambridge.

Thank Dr. Richard Parton and Dr. Kirsten Lillie of Department of Biochemistry, Oxford University, for providing the living cell images, and sharing their knowledge.

Thank Dr. Lei Yang for sharing his expertise in image processing and great help and suggestions on my thesis.

Thank Neil McCulloch and Adam W. Gripton for their great help in smoothing English for my thesis.

Thank my mates in Room 2.40 and many others in “Chaplaincy” for moral support over the past few years.

On a personal note, I would like to thank my wife, Ling Liu, for her love and encouragement.

## Table of Contents

Abstract .....	2
Acknowledgements .....	3
Chapter 1 .....	6
Introduction .....	6
1.1 Background .....	6
1.1.1 Wavefront Sensing and Zernike Filter .....	6
1.1.2 Image Processing and Zernike Filter .....	7
1.2 Layout and Original Contributions .....	8
Chapter 2 .....	9
Phasing of Segmented Telescopes .....	9
2.1 Segmented Telescopes .....	9
2.2 History of Segmented Mirror Telescopes .....	10
2.3 Phasing of Segmented Mirrors .....	12
2.3.1 Shack-Hartmann Wavefront Sensor .....	13
2.3.2 Curvature Sensor .....	15
2.3.3 Mach-Zehnder Interferometer .....	17
2.3.4 Pyramid Wavefront Sensor .....	19
2.3.5 ZEUS .....	22
2.4 Summary .....	24
Chapter 3 .....	26
Differential Zernike Phase Sensor for Phasing of Segmented Mirrors .....	26
3.1 Introduction .....	26
3.2 Mathematical Model of ZEUS .....	27
3.3 Differential Zernike Filter and DZEUS .....	31
3.4 Finite Aperture Size .....	33
3.5 Multiple Phase Jumps .....	36
3.6 Double Phase Jumps .....	39
3.7 Multiple Step Corrections .....	40
3.8 Two-Dimensional Simulations .....	41
3.9 Conclusions .....	45
Chapter 4 .....	46
Differential Zernike Feedback Phase Retrieval .....	46
4.1 Introduction .....	46
4.2 Theoretical Model .....	47
4.2.1 Zernike filter with Unknown Input Wavefront .....	47
4.2.2 Differential Zernike Filter .....	50
4.3 Wavefront Control Based on Gradient-Flow Optimization .....	50
4.3.1 Feedback-Controller Synthesis .....	50
4.3.2 System Performance Metric and Gradient-Flow Dynamics .....	52
4.3.3 Gradient-Flow Dynamics and Differential Zernike filter Synthesis .....	54
4.4 Differential Zernike Feedback Phasing Sensor .....	54
4.4.1 System Model of Differential Zernike Feedback Sensor .....	54
4.4.2 Small Phase Distortion Approximation .....	57
4.4.3 Simulation Results .....	59

4.5 Conclusions and Discussion.....	66
Chapter 5 .....	67
Automatic Particle Detection and Tracking in Living Cell.....	67
5.1 Method for Study of Biological Molecular Dynamics.....	67
5.2 Fundamental Problems of Automatic Tracking in Living Cell .....	68
5.3 Overview of Particle Tracking Methods.....	69
5.3.1 Image Pre-processing.....	70
5.3.2 Particle Detection .....	70
5.3.3 Particles Linking .....	71
5.3.4 Tracking Results Analysis.....	72
5.4 Summary .....	72
Chapter 6 .....	74
Zernike Filtering for the Application of Particle Detection and Tracking.....	74
6.1 Zernike Filter Pre-Processing .....	75
6.2 Image Pre-processing and Single Particle Tracking Procedure.....	78
6.3 Tests on Simulated Living Cell Images .....	79
6.3.1 Particle Model .....	80
6.3.2 Noise Model .....	81
6.3.3 Results .....	81
6.4 Tests on Real Living Cell Images.....	84
6.5 Multiple-Particle Tracking.....	86
6.5.1 The Tracking Algorithm .....	86
6.5.2 Tracking Results and Comparison with Manual Tracking.....	88
6.6 Conclusions and Discussion.....	91
Conclusions .....	95
Appendix A .....	97
Singular Value Decomposition.....	97
Appendix B .....	98
Atmospheric Turbulence Simulation .....	98
Appendix C .....	101
Single Particle Tracking Method.....	101
Publications.....	103
References .....	104

# Chapter 1

## Introduction

### 1.1 Background

#### *1.1.1 Wavefront Sensing and Zernike Filter*

Measuring the wavefront (optical phase) of a light beam is critical for assessing the quality of an optical system and optimizing its performance. When light passes through an imperfect optical component, aberrations – or wavefront errors – are generated. By measuring these errors with a wavefront sensor, they can be either corrected or minimized. Typical detection systems are ineffective in the analysis of such objects since the eye, CCD cameras, photomultipliers and other light detection devices are sensitive only to variations in intensity and not wavefront.

There are several types of wavefront sensors that are now commercially available. The most popular wavefront sensors are the Shack-Hartmann, curvature sensors, and multilateral shearing interferometers. Others, such as Pyramid sensor, March-Zehnder interferometer are also being investigated at present. Wavefront sensors can be found in a variety of applications, including performance assessment of aspheric lenses, characterization of DVD pick-up heads and the development of femtosecond lasers, compensation of atmospheric turbulence, and telescope adjustment.

The Zernike filter, suggested by Zernike (for which he was awarded the Nobel Prize in 1953), is one among the wavefront sensing techniques [31]. It is known as Phase

Contrast Technique, which converts phase modulation caused by the object into amplitude modulation. In the case of weak phase objects, this conversion is performed such that phase variations map linearly to amplitude variation in the image field, thus rendering an image depicting the variation of optical thickness of a phase object.

Zernike filter has been applied for static atmospheric turbulence compensation [34, 35], and phasing of segmented astronomy telescopes [15]. Segmented primary mirror alignment, known as phasing, is one of the critical tasks of telescope adjustment. For accurate alignment of the segmented primary mirror, it is necessary to measure the extent of wavefront aberration of the light due to imperfectly aligned primary mirror surface. One of the objectives in this thesis is to utilise Zernike filter in wavefront sensing technique for phasing of segmented telescopes.

### ***1.1.2 Image Processing and Zernike Filter***

Automated particle tracking and analysis in image sequences is one of the major fields in digital image analysis research. There are many applications in video surveillance, multimedia services, automated vehicle guidance and driver assistance, remote sensing and meteorology, and medical imaging. It is also a very important theme in molecular biology. By their very nature, biomolecular systems are dynamic, and it is one of the major challenges of biomedical research and pharmaceutical industries to unveil the spatial and temporal relationships of these complex systems. Results in this area can be expected to have significant social and economic impact in the near future, as they can improve human health and well-being. Studies into biomolecular dynamics generate ever-increasing amounts of image data. To be able to handle these data and to fully exploit them for describing biological processes on a quantitative level and building accurate mathematical models of dynamic structures, computerized motion analysis is becoming a necessity. Over the past decade, a number of image analysis techniques have been developed in support of such studies.

Achieving robustness and high accuracy in particle tracking and motion analysis in images obtained by light microscopy is hampered by three factors: the limited spatial resolution of the microscope; low signal-to-noise ratio as a result of the quantum nature

of light; and large variability of biological image data. All these factors put high demands on the design of automated image analysis techniques.

Zernike filter can be applied to image processing to solve one of the essential problems - low signal-to-noise ratio - within particle tracking by adapting it into a feedback filtering system. We use the Zernike filter in this application because in Zernike filter feedback system, static component (noisy background) can be suppressed, and hence increasing the signal-to-noise ratio.

## **1.2 Layout and Original Contributions**

This thesis can be divided into two separate parts. In the first part, which includes Chapter 2, 3, and 4, we study Zernike filter based segmented mirror phasing techniques. In Chapter 2, we present a review on existing segmented mirror phasing techniques. My original work in the first part of the thesis is presented in Chapter 3 and 4. Our investigations are by means of combined computational simulations and theoretical analysis mainly through Fourier transform. In Chapter 3, the differential Zernike filter is introduced and its application in segmented mirror phasing is studied both analytically and numerically. In Chapter 4, the differential Zernike filter is further integrated into an adaptive optical system, which is realized by combining the differential Zernike filter with a two-dimensional feedback.

In the second part, which includes chapter 5, and 6, we study the image pre-processing, particle detection and tracking method especially in the cell biology application. Chapter 5 provides a review on the particle tracking in living cells. My original work for the second part of the thesis is presented in Chapter 6. In Chapter 6, we introduce the concept of Zernike filtering, which originate from that of Zernike feedback system, is used along with an existing single particle tracking algorithm to achieve better signal-to-noise ratio. Finally, we conduct automatic tracking on a real living cell image sequence, as an example of high particle density case, by an existing efficient multiple-particle tracking algorithm with a few modifications.



## Chapter 2

### Phasing of Segmented Telescopes

#### 2.1 Segmented Telescopes

The size of the primary mirror is critical to the observation capability of a telescope. Many telescopes are limited by the amount of light collected from the astronomical object. Angular resolution is also often a critical factor to understanding the nature of the astronomical targets. The angular resolution of a telescope mirror at light wavelength  $\lambda$  is determined by the relation  $\lambda/D$  (Rayleigh's Criterion). Thus, a telescope with a larger primary mirror will have a higher resolution, allowing it to accurately image smaller details. The needs of the astronomical community have led to the development of telescopes with mirrors of ever-greater diameter.

Two decades ago, nearly all telescopes contained monolithic telescope mirrors, which are mirrors comprised of a single piece of glass. Building a giant telescope from a monolithic mirror has many difficulties. Some difficulties typically grow rapidly with the increasing size, and quickly make monolithic mirrors impractical. These difficulties are:

- A) Reduced availability of mirror blank material
- B) Large optical deflections as a result of passive support of mirror
- C) High risk of mirror breakage from mishandling
- D) Larger deformations from thermal changes due to larger mirrors
- E) Large tool costs for all parts (fabrication and handling)
- F) Shipping being difficult

As the diameter of a monolithic mirror is increased, the thickness of the mirror must also be increased. Eventually, the sheer size and weight of the glass required makes this impractical for use in ground-based telescopes above a certain diameter. Similarly, the issues also make them impractical for use in larger space-based platforms due to spacecraft weight considerations.

This limitation has been overcome with the development of segmented telescope mirrors, a concept originally proposed by Nelson *et al.* for use in the Keck Telescopes [1]. In a segmented telescope mirror, many smaller hexagonal mirror segments are placed side by-side forming a single, continuous main mirror. Because of their smaller individual size, the individual segments do not need to be as thick as a single monolithic mirror with the same total diameter. With the weight limitation removed, the construction of extremely large telescopes becomes possible.

## **2.2 History of Segmented Mirror Telescopes**

People have divided regions into segments for ages, ranging from bathroom tiles to modern segmented mirror telescopes. Even the application of segmentation to optics is old. The first recorded use of segmented mirrors was by Archimedes, who in 212BC had an array of mirrors focused on attacking Roman navy in order to defend Syracuse. More recently, Horn d'Arturo in Italy made a 1.5 m segmented mirror in 1932. It was only used vertically, and was not actively controlled. In the 1970's, Pierre Connes in France made a 4.2 m segmented mirror telescope for infrared astronomy. It was fully steerable, and active. Unfortunately, the optical quality was too low to be very useful for astronomic observation.

Another type of segmented mirror telescope (actually multiple telescopes on a common mount) was developed in the 1970's and completed early in the 1980's. This was called Multiple Mirror Telescope (MMT), and was built in southern Arizona. The telescope was made of six 1.8 m primary mirrors, each axis-symmetric. Although this telescope worked, it suffered from a number of problems, and was not viewed as very successful. In the late 1970's a very ambitious project to build a 10-m diameter segmented mirror telescope began, called the Keck Observatory. The geometry of the segmented primary mirror of the Keck telescope is shown in Fig. 2.1. The segmented

mirror telescope was very successful and due to its success, funds were acquired to make a second Keck telescope, and it was positioned to allow the two Keck telescopes to be used for interferometry as well as individual telescope observing.

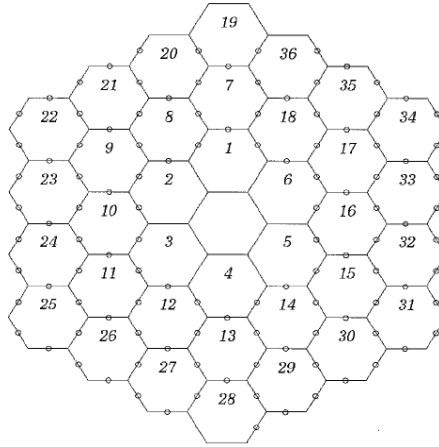


Figure 2.1 Geometry of the primary mirror of the Keck telescopes. Each segment is 0.9 m on a side, showing the 78 circular subapertures that sample the intersegment edges in the phasing procedure. The subapertures are 12 cm in diameter.

Since the existence of Keck, many other segmented telescopes have been proposed and built. The Southern African Large Telescope (SALT) is a ~10 metre diameter optical telescope, located in the semi-desert region of the Karoo, South Africa. Similar to the Keck Telescopes, the primary mirror is composed of an array of mirrors designed to act as a single larger mirror; however, the SALT mirrors produce a spherical primary, rather than the parabolic shape associated with a classical Cassegrain telescope. Each SALT mirror is a 1-meter hexagon, and the array of 91 identical mirrors produces a hexagonal-shaped primary with a size of 11 x 9.8 meters. The Gran Telescopio Canarias (GTC) ("Great Telescope Canary Islands"), sometimes called GranTeCan, is a 10.4m reflecting telescope and is undertaking commissioning observations at the Observatorio del Roque de los Muchachos on the island of La Palma, Spain. The telescope is sited on a volcanic peak 2,400 metres above sea level. The GTC started preliminary observing on 13 July 2007 following an opening ceremony using 12 segments of its primary mirror later to be increased to a total of 36 hexagonal segments fully controlled by an active optics control system.

There are also some other segmented telescopes under plan or construction. The Thirty meter telescope (TMT) (formerly called the California Extremely Large Telescope (CELT)) is a future large segmented-mirror optical and infrared extremely large telescope. Completion is scheduled for November 2014. Its primary mirror is designed to consist of 492 hexagonal segments each about  $\sim 1.4$  m in diameter. The Overwhelmingly Large Telescope (OWL) is a conceptual design by the European Southern Observatory (ESO) organization for an extremely large telescope, which was intended to have a single aperture of 100 meters in diameter, but was later scaled down to a 60 meter diameter telescope. Because of the complexity and cost of building a telescope of this unprecedented size, ESO has decided to focus on the less ambitious 42 meter diameter European Extremely Large Telescope instead.

### 2.3 Phasing of Segmented Mirrors

Although segmented mirrors are a promising option by comparison with monolithic primary mirror, they also bring about their own problems. The most well-known of these problems is achieving a smooth continuous mirror surface, a process known as phasing. A properly phased telescope will have a resolution comparable to the total diameter of the entire segmented primary mirror. On the other hand, a pre-phasing telescope will have very poor resolution, which is limited by the diameter of an individual segment. The importance of the phasing of the segment mirrors in a telescope has been demonstrated in several publications [2-4].

The analysis in this thesis is restricted to the correction of the piston, tip and tilt errors of the segments. Piston errors represent segment vertical misalignment in segmented mirror. Tip-tilt errors represent the segment turned up or down at the inter-segment edge. One important process in phasing a segmented mirror telescope involves the vertical displacements (piston error) between adjacent segments. If the piston errors between the segments are greater than approximately  $\lambda/20$  (where  $\lambda$  is the wavelength of light), then the effective diameter of the telescope (D) is equal to the diameter of a single mirror segment. Only when the piston errors are reduced below  $\lambda/20$  is D given by the total diameter of the entire segmented mirror. In other words, the

full angular resolution of the entire segmented primary mirror is obtained only when the piston errors between the segments are smaller than  $\lambda/20$ .

Clearly, it is important that the piston errors are reduced to less than  $\lambda/20$  if the optimal resolution offered by the segmented mirror is to be achieved. While each mirror segment may be raised and lowered independently of the others by a set of actuators underneath, the issue of accurately determining the piston errors between the segments to a high enough precision is critical. This makes it very important to develop what are known as phasing algorithms: operations used to detect and correct the discontinuities between mirror segments through the analysis of various diffraction phenomena along the inter-segment edges.

Three hardware systems are required for the active segment control: segment edge sensors which provide real time information about the relative segment displacements, segment actuators which compensate these displacements, and a phasing camera which measure the phasing errors optically [5, 52, 53].

The optical phasing camera is used at the beginning of each night before observation to measure the phasing errors. These measurements can be then used to control the segment actuators. The two steps can be repeated for several times to achieve a desired accuracy. During the operation of the telescope the phasing corrections are based on signals from edge sensors at intersegment borders. The reading of the edge sensors at the beginning of the night achieved by optical phasing is used as a reference.

There are several existing optical phasing techniques reported by other researchers in the past years, including a modified Shack-Hartmann wavefront sensor [6-8], curvature wavefront sensor [9, 10], Mach-Zehnder interferometer [11, 12, 51], Pyramid wavefront sensor [13, 14], and ZEUS (Zernike Unit for Segment phasing) [15].

### ***2.3.1 Shack-Hartmann Wavefront Sensor***

The Shack-Hartmann wavefront sensor is a simple and elegant means for measuring the shape of a wavefront. This technique has found application to a wide variety of applications [16, 47]. Among them, a modified Shack-Hartmann wavefront sensor [6] is used for detection of phasing errors in Keck Telescopes.

The principle of traditional Shack-Hartmann wavefront sensor is shown in Fig. 2.2. An image of the exit pupil is projected onto a lenslet array – a collection of small identical lenses. Each lens takes a small part of the aperture, called sub-pupil, and forms an image of the source. All images are formed on the same detector, typically a CCD. When an incoming wave-front is plane, all images are located in a regular grid defined by the lenslet array geometry. As soon as the wavefront is distorted, the images become displaced from their nominal positions. Displacements of image centroids in two orthogonal directions are proportional to the average wavefront slopes in either direction over the subapertures. Thus, a Shack-Hartmann wavefront sensor measures the wavefront slopes.

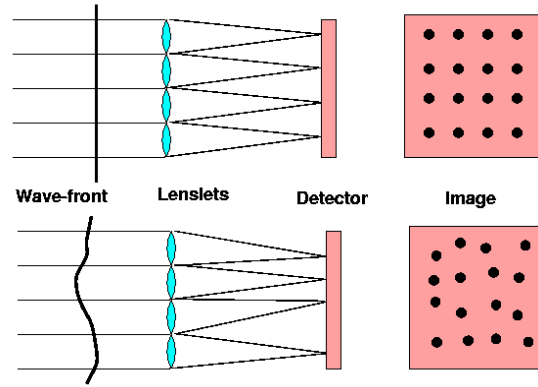


Figure 2.2 Schematic of the traditional Shack-Hartmann Wavefront sensor.

In Keck telescopes, the key element of Shack-Hartmann phasing camera is an array of  $2mm \times 3mm$  prisms, which replace the usual lenslet array in a traditional Shack-Hartmann wavefront sensor. This prism array is preceded by a mask at the position of the exit pupil. The mask, at a scale of  $1/200$  of the primary mirror, defines small circular subapertures at the centre of each intersegment edge. The size of the subapertures is chosen to be significantly smaller than the atmospheric coherence diameter, to ensure that the results will be insensitive to atmospheric turbulence. The atmospheric coherence diameter corresponds to the length-scale over which the turbulence becomes significant (10-20 cm at visible wavelengths at good observatories).

There are two algorithms, the narrowband and broadband algorithms, which corresponds to the incoming light sources being monochromatic and of a finite bandwidth respectively [8, 48-50] along with Shack-Hartmann phasing camera. For

narrowband algorithm, phase differences between segments manifest themselves in the resulting diffraction patterns on the detector, and these phase relationships can be extracted by cross-correlation. Because the diffraction patterns are a periodic function of the phase difference between the two segments, the capture range of the narrowband algorithm (i.e. the maximum piston error which can be reliably detected) should be about  $\lambda/4$ . The broadband algorithm is similar to the narrowband and uses identical hardware, but it exploits the finite bandwidth of the filters which define the wavelength of the starlight. The signal is the degree of coherence of the sub-image, and the relevant scale is not the wavelength but the coherence length of the filter. This is not only in general much larger than the wavelength, but it can be tuned to the conditions at hand if one has an a priori estimate of the phase errors involved. As a result this technique has an enormously large capture range and also dynamic range (This is roughly the coherent length, for a specific example. Consider a filter with a bandwidth of 10 nm and a central wavelength of 891 nm, corresponding to a coherence length of  $40\ \mu\text{m}$ ).

The mask-pupil registration is critical to this modified Shack-Hartmann scheme as the subapertures must be aligned accurately with respect to the intersegment edges, which in turn requires precise hardware positioning of lenslets with respect to a reimaged telescope pupil. As the number of segments grows, this requirement becomes increasingly critical, imposing extremely tight specifications on the pupil reimaging optics, particularly in terms of distortion [12]. As a result, the phasing camera designed for Keck may not be so easy to implement for a telescope consisting of more than 600 segments.

### 2.3.2 Curvature Sensor

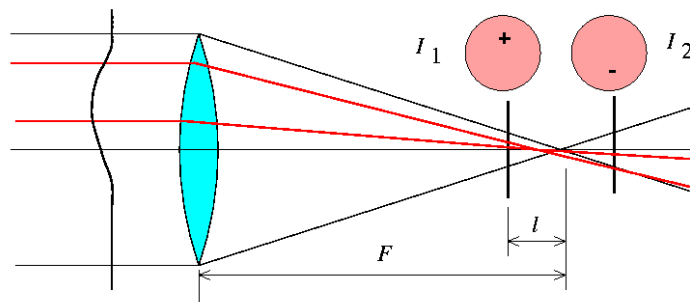


Figure 2.3 A schematic of a curvature wavefront sensor.

The curvature wavefront sensing was developed by F. Roddier since 1988 [17-19]. It takes an entirely different approach than Shack-Hartmann wavefront sensors. The curvature sensor measures an “image” at a location between the pupil plane and the image plane. If this image is before focus, it is called the intrafocal image; the image after focus is the extrafocal image, as shown in Fig. 2.3. The intrafocal image will be brighter in regions with positive curvature and darker in regions with negative curvature. The intensity pattern of the extrafocal image will be reversed with respect to that of the intrafocal image. In principle, only one out of focus image is needed to measure wavefront curvature. However, using both the intrafocal and extrafocal images makes a curvature system work better for several reasons: automatic compensation of systematic errors - variation in quantum efficiency, electronic gain, etc., and compensation of atmospheric scintillation.

Using the vector  $\vec{r}$  for the location  $x, y$  in a  $z$ -plane,  $I_1$  and  $I_2$  for the intrafocal and extrafocal images, the signal  $S$  can be constructed from the sum and difference of the intrafocal and extrafocal images:

$$S(\vec{r}) = \frac{I_2(\vec{r}) - I_1(-\vec{r})}{I_2(\vec{r}) + I_1(-\vec{r})}. \quad (2.1)$$

To ensure that  $S(\vec{r})$  carries useful information on the segment phases, two conditions on the wavelength  $\lambda$  must be satisfied [9]:

- (a) The scale of diffraction effects (associated with primary mirror segments) in the image plane should be small compared to the diameter  $d$  of a segment mapped onto the image plane,  $\frac{f\lambda}{d} \ll \frac{dl}{f}$ . Here,  $f$  is the focal length of the telescope and  $l$  is the defocus distance.
- (b) Diffraction effects associated with the segments should predominate over those associated with the atmosphere,  $r_0(\lambda) \gg d$ .

To get some insight into this method, consider an out-of-focus image of a Keck mirror, as shown in Fig. 2.4, perfectly aligned except for segment 13, which has a piston error of  $\lambda/8$  [20]. The dominant feature in this image is well-localized at a position



on the detector which has an obvious approximate correspondence to the location of the segment in question in the pupil. For piston errors of this size or smaller, the strength of the feature in the difference image will vary monotonically with the piston error. It follows that multiple piston errors can therefore be extracted from the difference image by straightforward cross-correlation techniques. Because the diffraction pattern in Fig. 2.4 spills over the boundaries corresponding to segment 13 and because the linearization effected by the restriction to small piston errors is only approximate, the curvature phasing algorithm does not converge in one step, but rather is an iterative procedure, requiring multiple (typically 5-6) exposures.

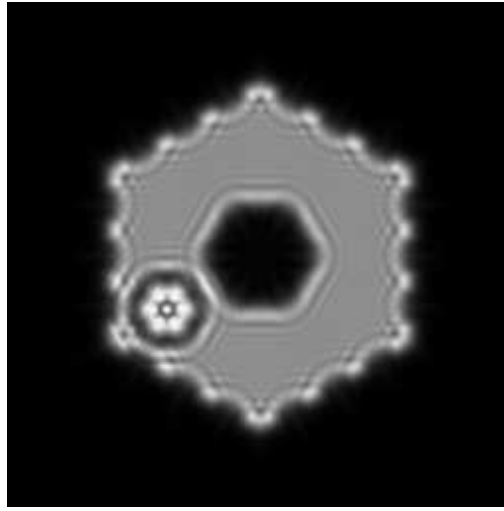


Figure 2.4 Numerically generated out-of-focus image of the Keck telescope with segment 13 pistoned by  $\lambda/8$ . Note that the resulting diffraction effects are well-localized at a position that has an obvious correspondence with the position of the segment in the pupil.

Curvature wavefront sensing has the immediate advantage that no special purpose hardware is required, only a detector is needed. The precise mask-pupil registration is avoided, and it utilizes a larger fraction of the segment surface by comparison with Shack-Hartmann techniques.

### 2.3.3 *Mach-Zehnder Interferometer*

The Mach-Zehnder interferometer, developed a century ago to measure the refractive index variation in a suppressible gas flow, is an example of a classical optical system

which finds uses in various applications. A schematic illustration of the Mach-Zehnder sensor is shown in Fig. 2.5. The beam is split into two arms of the interferometer, and a pinhole, placed in the focal plane of one arm, acts as a spatial filter providing a reference wave coherent with the beam in another arm. The two beams are recombined and form two complementary interference patterns recorded by two imaging detectors. The proposed sensor departs from this classical scheme by the size of the pinhole. While the classical versions employ a pinhole smaller than the Airy disk, producing a perfectly spherical reference beam, the version that we propose for segment phasing uses a much larger spatial filter whose diameter is approximately equal to that of the seeing disk. The seeing disk, in astronomy, is a reference to the best possible angular resolution which can be achieved by an optical telescope, which is viewing the celestial sphere from within an atmosphere. The reference wave is now a low-pass filtered version of the original wave front, which, when interfering with the latter, produces an intensity distribution conveying information about only the high-frequency wave-front aberrations. While the atmospheric turbulence has a power spectral density that is dominated by low-frequency errors and that falls off rapidly toward higher frequencies according to the Kolmogorov law [42], the power spectral density representing intersegment phase steps has strong high-frequency components. Eliminating low-frequency wave-front errors will therefore make piston-induced errors dominant. Furthermore, since the remaining aberrations related to atmospheric turbulence are smaller than 1 rad, they will average out in a long exposure image.

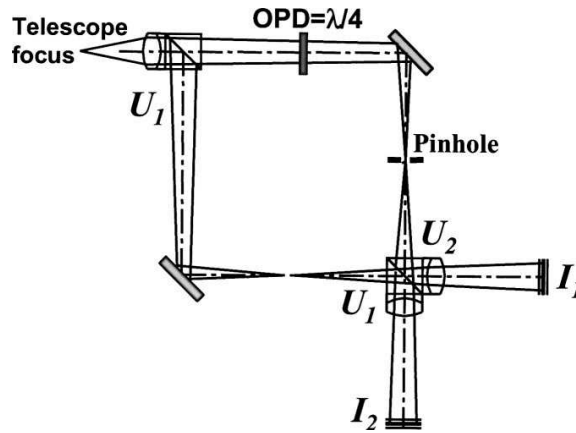


Figure 2.5 Schematic representation of the Mach-Zehnder phasing sensor.

We denote by  $I_1$  and  $I_2$  the intensity in the two arms of the interferometer measured in a pupil plane. If the OPD is equal to  $\lambda/4$ , the one dimensional signal of Mach-Zehnder interferometer can be written as

$$S(x) = I_1(x) - I_2(x) = \sin(\Delta\phi)\text{sign}(x)[1 - \Phi(b|x|)], \quad (2.2)$$

where  $x$  is the one-dimensional spatial vector,  $\Delta\phi = \phi_2 - \phi_1$  is a phase “jump”, and  $b$  is related with the pinhole size  $a$  by the relation  $b = 0.6\pi a / \lambda$ . The function  $\Phi(bx)$  is the error function:

$$\Phi(bx) = \frac{2b}{\sqrt{\pi}} \int_0^x \exp(-b^2 x'^2) dx'. \quad (2.3)$$

The expression for the signal is factorized: While the amplitude of the signal is a sine function of the relative piston, its width is inversely proportional to the width of the pinhole.

An algorithm for phase reconstruction from the signal  $S(x)$  can use the maximum of the signal amplitude, or the difference between peak and valley, or the difference between the integrals of the “positive” and “negative” parts of the signal. Using the latter option, the calibration function can be defined as

$$K(\Delta\phi) = \int_0^{x_0} S(x)dx - \int_{-x_0}^0 S(x)dx = \sin(\Delta\phi) \{ 2 \int_0^{x_0} [1 - \Phi(bx)]dx \}, \quad (2.4)$$

where the integration area  $x_0$  is a free parameter that has to be optimized.

The integral criterion of Eq. 2.4 has been used in a close-loop phasing algorithm [12]. As Eq. 2.4 is a sine function of  $\Delta\phi$ , the range of the measurable phase difference is limited to  $[-\pi/2, \pi/2]$ . However, this capture range can be enlarged up to  $\pm \lambda/2$  by shifting the pinhole in the direction orthogonal to the considered intersegment border. This shift can be realised by adding a known tilt to the incoming wavefront [12].

### 2.3.4 Pyramid Wavefront Sensor

The concept of the pyramid wavefront sensor is based on a modification of the Foucault knife-edge test used in optics to evaluate qualitatively the aberrations of an optical system. Pyramid wavefront sensing was first proposed by Ragazzoni in 1996

[21]. Like a Shack-Hartmann sensor, it measures the first derivative of the wavefront, the wavefront slope or gradient.

A pyramid lens with four equal faces is placed with its vertex on the nominal focus point of the optical system. The four faces deflect the beam in four different directions, depending on which face of the prism gets hit by the incoming ray. A field lens is then used to re-image the pupil of the telescope. In the pupil plane a detector is used to measure the individual signals of the four faces. The principle is illustrated in Fig. 2.6.

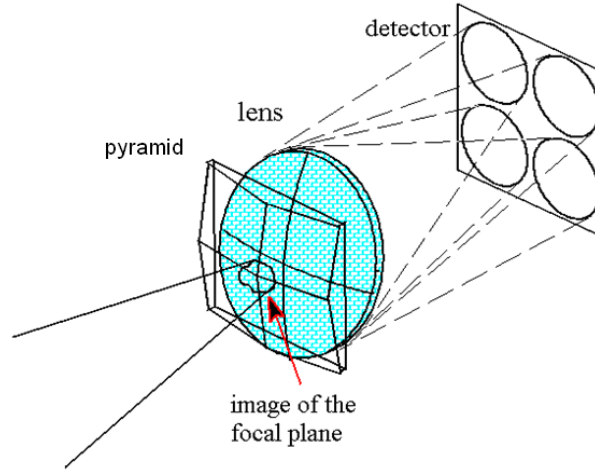


Figure 2.6 Principle of the pyramid wavefront sensing. The pyramid lens is oscillating. When the incoming light reaches one of the four faces, it will be deflected in slightly different directions, and forming four pupil images on the detector surface. If the incoming light suffers aberrations, the four pupil images are no longer equal and from the relative point-to-point intensity differences the local gradient can be computed.

In this configuration a ray of the incoming beam with wavefront error  $\psi(\vec{r})$ , originating from a generic point  $\vec{r} = (x, y)$  on the pupil plane, is aberrated and reaches the pyramid displaced by a vector  $\vec{p}$  from the vertex. The amount of displacement is:

$$\vec{p} = f \frac{\partial \psi(\vec{r})}{\partial \vec{r}}, \quad (2.5)$$

where  $f$  is the effective telescope focal length.

Hence one face will refract the ray and only in the corresponding pupil the region conjugated to the point  $\vec{r}$  in the pupil plane will be bright. The other three pupils will

show a dark region in the same point. Although it is not possible to obtain the value of the aberration, its sign can be retrieved.

Ragazzoni proposed to oscillate the pyramid (parallel to its base plane in " $dx$ " and " $dy$ " direction) to allow all the aberrated rays to sweep over the four faces. If the introduced modulation  $\delta\theta_x$  in the  $x$  and  $\delta\theta_y$  in  $y$  direction satisfies the requirement  $\delta\theta > \left| \frac{\partial\psi(\vec{r})}{\partial\vec{r}} \right|$ , each pupil will receive a particular intensity of illumination that will be proportional to the displacement of the rays with respect to the pyramid vertex.

After an integer number of oscillation cycles, the four pupil signals  $I_1$ ,  $I_2$ ,  $I_3$  and  $I_4$  are combined and normalized by the sum. Hence it is possible to retrieve the first derivatives or slopes  $\frac{\partial\psi(x,y)}{\partial x}$  and  $\frac{\partial\psi(x,y)}{\partial y}$  of the wavefront along two orthogonal axes:

$$\frac{\partial\psi(x,y)}{\partial x} = \delta\theta_x \frac{(I_1 + I_4) - (I_2 + I_3)}{I_1 + I_2 + I_3 + I_4}, \quad (2.6)$$

and

$$\frac{\partial\psi(x,y)}{\partial y} = \delta\theta_y \frac{(I_1 + I_2) - (I_3 + I_4)}{I_1 + I_2 + I_3 + I_4}. \quad (2.7)$$

The pyramid wavefront sensor shows some advantages with respect to a Shack-Hartmann sensor. The gain is variable by adjusting the amount of the movement or oscillation of the pyramid in  $x$  and  $y$ . The amplitude can be increased when the image tilt is too large (for example during closing the loop), or to maintain an as high as possible signal to noise ratio during the measurements.

The pyramid wavefront sensor was shown to be able to detect signals that are due to phasing error among segments of a segmented mirror for the first time in 2001 [22]. Numerical simulations performed demonstrated that the sensor can be used in an iterative control loop to phase a segmented mirror by using monochromatic light. It was pointed out in connection with those simulations that the sensor can simultaneously sense and control segment differential piston plus the tip and tilt of each individual

segment without changing anything in the wavefront sensor configuration. A first experimental investigation of the pyramid wavefront sensor as phasing sensor was reported in 2003 [13]. The measured signal was found to be in agreement with the numerical simulations. In a more recent experiment [14] the alignment of a segmented mirror having three degrees of freedom per segment, i.e., differential piston, tip, and tilt, was studied. The closed-loop procedure reached a typical average wavefront residual error after loop convergence of 10 and 15 nm for piston and for tip and tilt, respectively. The results show the ability of the pyramid wavefront sensor to phase and align a segmented mirror in terms of piston and tip and tilt.

### 2.3.5 ZEUS

The ZEUS (Zernike Unit for Segment Phasing) concept is developed in the context of the ESO (European Organisation for Astronomical Research in the Southern Hemisphere)-led Active Phasing Experiment (APE). It has its origin in the Mach-Zehnder (MZ) phasing sensor concept. In ZEUS, the Mach-Zehnder interferometer is replaced by a simple phase mask. ZEUS is almost analogous to the Mach-Zehnder concept physically, and it shares most of the performance characteristics of the Mach-Zehnder while avoiding the delicacy of the interferometric setup.

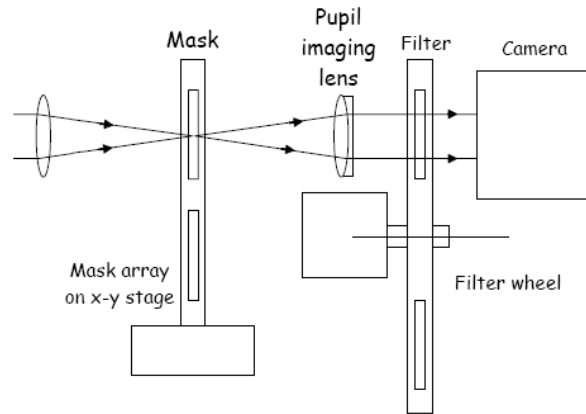


Figure 2.7 Optical layout of the ZEUS system.

Figure 2.7 shows schematically the optical layout of ZEUS. A Zernike phase plate, which is made of a transmissive mask with a diameter of the size of the full width at half maximum of the seeing disk (around 0.6 arcsec) etched onto a glass plate, is placed

at the common focal plane of a 4f system. The thickness of the mask corresponds to an optical path difference of  $\pi/4 - \pi/2$ . This mask replaces all the Mach-Zehnder interferometer optics. It must be mounted on a selection mechanism in order to provide a straight through reference image and a choice between two or more masks of different thickness and/or diameters to account for different operational conditions (seeing, coarse or fine phasing, etc). A lens projects the pupil onto the camera, producing images with the phasing errors contained in. A filter wheel provides band selection; narrow-band and multi-band operation will be used for coarse phasing, and broad-band operation will be used for ultimate performance fine phasing.

The signal image ( $I_S$ ) is normalized using three measured images, an image ( $I$ ) with the mask in place, the reference image ( $I_R$ ) obtained by removing the mask, and a dark frame ( $I_D$ ) obtained by blocking the light path:  $I_S = (I - I_R)/(I_R - I_D)$ . The signal typically shows a sharp positive-negative double peak, as seen in Fig. 2.8.

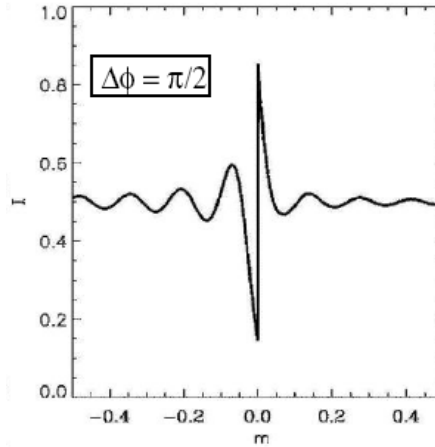


Figure 2.8 ZEUS signal profiles in the absence of atmosphere and for piston phase values of  $\pi/2$ . The mask thickness corresponds to a phase difference of  $\theta = \pi/4$ .

Its half peak-to-peak (PtP) amplitude, which can be determined by searching for positive and negative peaks near the segment edge, can be used to determine the inter-segment phase step ( $\Delta\phi$ ) according to approximation:

$$S = PtP / 2 = A \sin(\Delta\phi), \quad (2.8)$$

where A is the calibration constant. The value of piston error in the wavefront can be retrieved by measuring the PtP value, and their relation can be given by:

$$p = \Delta\phi\lambda/(2\pi) = \sin^{-1}(S/A)\lambda/(2\pi). \quad (2.9)$$

In the absence of atmospheric turbulence,  $A$  depends only on the mask thickness. The calibration curve has a period of  $2\pi$ , corresponding to a phase step of  $\lambda$  on the wavefront ( $\lambda/2$  on the mirror surface) and an un-ambiguity range of  $\pm\pi/2$ . Note that when operating in closed loop, the single-wavelength capture range, within which the closed-loop feedback system will converge towards  $p=0$ , is  $\pm\pi$ .

Results of simulations and lab experiment have shown that the phase retrieval algorithm (Eq. 2.9) can be used for system optimization purposes. Precision estimates indicate that piston errors of 3nm can be measured in fine-phasing mode by observing a 10th magnitude guide star through a broad-band filter with 100s exposures [15].

## 2.4 Summary

In this chapter, we have presented various techniques that have been developed and proposed for optical phasing of the segmented primary mirror. Without optical phasing, the misalignments of the segments can be of the order of several micrometers [20]. The basic principle for most of these techniques is a modification of the wave front reflected by the mirror surface in such a way that the amplitude or its distribution pattern of the detected wave conveys the information about the phase discontinuities or the derivatives of the wavefront. In Shack-Hartmann sensor, different phase steps give different diffraction patterns. In curvature sensor, the phase discontinuities are retrieved by measuring the difference in intensity between images obtained equal distances before and after the telescope focus. In pyramid wavefront sensor, a refractive element (the pyramid) is used to produce four images of the entrance pupil, and the phasing errors are derived by an algorithm based on these four images. In the Mach-Zehnder sensor the phase discontinuities are revealed by the difference of the two interferograms of the Mach-Zehnder interferometer. ZEUS in many ways shares the performance characteristics of the Mach-Zehnder but avoids the delicacy of the interferometric setup by using Zernike phase mask. In most of these methods, the piston and tip/tilt can be retrieved from the intensity pattern in the output signal. However, the output signals attributed to different methods contain different amounts of information related to



segment errors, therefore the measurement precision varies for different methods. Most of these sensors are compared with each other within the Active Phasing Experiment (APE) project [27, 28]. We also note that when the phasing errors exceeds  $\pm \pi/2$ , most of the narrow-band phasing techniques suffer from a  $2\pi$  ambiguity. This can be solved with multi-wavelength techniques [8, 23-26].

In the next chapter of this thesis, we discuss a new segmented mirror phasing technique (DZEUS) which has developed from the concept of ZEUS. In chapter 4, we study differential Zernike filter in a feedback system, in which the phase is automatically retrieved without any algorithmic computation.

## Chapter 3

### Differential Zernike Phase Sensor for Phasing of Segmented Mirrors

#### 3.1 Introduction

As we discussed in the last chapter, the requirement for precise hardware positioning of the reimaged pupil may be avoided in phasing sensors that are based on pupil plane detection. Four phasing sensors are currently studied for this purpose, they are, curvature sensor, pyramid sensor, Mach-Zehnder interferometer, and Zernike unit for segment phasing (ZEUS). The basic idea for all these four techniques is that the intensity of the outgoing wave conveys the information about the phase discontinuities of the segmented mirror. In curvature sensor, the phase discontinuities is retrieved by measuring the difference in intensity between images obtained equal distances before and after the telescope focus. In pyramid wavefront sensor, a refractive element (the pyramid) is used to produce four images of the entrance pupil, and the phasing errors are derived by an algorithm based on these four images. In the Mach-Zehnder sensor the phase discontinuities are revealed by the difference of the two interferograms of the Mach-Zehnder interferometer. ZEUS in many ways shares the performance characteristics of the Mach-Zehnder but avoids the delicacy of the interferometric setup by using Zernike phase mask. In ZEUS, the peak-to-peak (PtP) value of a localised intensity variation in the output has approximately a linear relation to sine function of a phase jump (piston) of  $\phi$  in the input light [15], i.e.,  $PtP = A \sin \phi$ , where  $A$  is the

calibration constant. The PtP value is therefore used to determine the corresponding piston error. The measurement accuracy in ZEUS is mainly determined by how good the approximation of the linear relation is.

As we know, the output intensity of ZEUS can be seen as two parts: symmetrical part and anti-symmetrical part [15]. The anti-symmetrical part is the main carrier of the phasing information (PtP). The symmetrical is usually seen as a source of pollution of the signal.

In this chapter we present a differential Zernike filter based sensor for phasing of a segmented mirror. Differential Zernike filter is developed based on the traditional Zernike filter, and it was originally adopted in wavefront sensing for improved image contrast over conventional Zernike filter [29, 30]. This system is realized by replacing Zernike filter with differential Zernike filter in the ZEUS scheme, referred to as DZEUS. In section 3.3, we present in one-dimensional case the analytic expression for the PtP values of output intensity modulations against the input phase jumps in DZEUS. We show that the DZEUS gives rise to a better linear relation, compared to ZEUS, because the differential algorithm used in DZEUS remove the symmetrical (pollution) term that exist in the PtP expression of ZEUS. This makes DZEUS a better phase retrieval algorithm. In order to further improve the phasing accuracy, we put forward a multiple step correction approach which can further reduce the phase errors by iterations. In Section 3.7, we extend the one-dimensional analysis to two-dimension and study the performance of DZEUS using numerically generated segmented mirrors. We show that using the multiple step correction approach, the phasing error is reduced exponentially on the increase of iteration number. We further compare the characteristics between DZEUS and ZEUS and conclude that the former performs consistently better. Finally, we show that DZEUS is robust with respect to the atmospheric turbulence.

### **3.2 Mathematical Model of ZEUS**

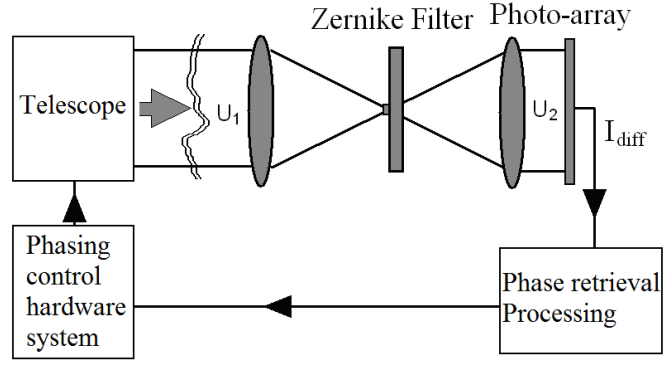


Figure 3.1 Schematic of the ZEUS phasing camera integrated with Segmented Telescope.

The main component of ZEUS is a traditional Zernike filter [31-33]. A traditional Zernike filter consists of two lenses with a phase-changing plate (Zernike phase plate) placed in the lenses' common focal plane. The phase plate has a small circular region (a dot) in the middle that introduces a phase shift of usually  $\pi/2$  rad into the focused wave; the size of the circular region typically has the diffraction-limited radius of a focused, undistorted input wave. In practice, a narrow band filter can be applied before the Zernike filter, and we assume the incoming light is monochromatic. The phase-shifted wave-front then constitutes an in-line reference with which the radiation scattered from the phase object interferes and the resulting phase contrast can be observed. The output intensity distribution of Zernike filter can be measured by a photo detector. A schematic of Zernike filter in the ZEUS configuration is shown in Fig. 3.1. If we denote by  $U_1(\vec{r})$  and  $U_2(\vec{r})$  the complex amplitudes at the input and output pupil plane respectively, where  $\vec{r}$  is the two-dimensional position vector, the complex amplitude in the common focal plane of the Zernike filter can be written as a Fourier transform of the input wave:

$$u_1(\vec{k}) = \int U_1(\vec{r}) \exp(i\vec{k} \cdot \vec{r}) d\vec{r}, \quad (3.1)$$

where  $\vec{k}$  is a two-dimensional wave number in the focal plane, while  $\vec{k} = \frac{2\pi}{\lambda f} \vec{\xi}$ , and

$\vec{\xi}$  is the two-dimensional position vector in the focal plane, and  $\lambda$  is the central operating wavelength of the telescope, and  $f$  is the focal length of the lens. The

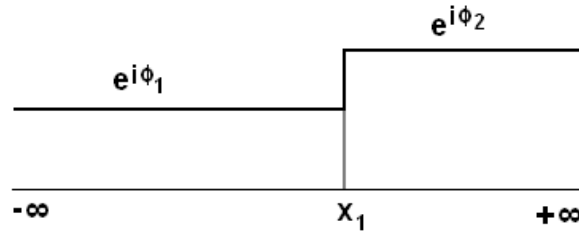
complex amplitude of the wave in the output pupil of the Zernike filter after the phase filter is:

$$U_2(\vec{r}, \theta) = \frac{1}{2\pi} \int u_1(\vec{k}) t(\vec{k}, \theta) \exp(-i\vec{k} \cdot \vec{r}) d\vec{k}, \quad (3.2)$$

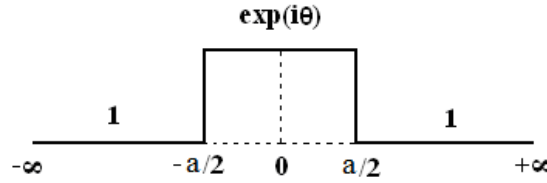
where  $t(\vec{k})$  is the filter function of Zernike phase plate:

$$t(\vec{k}, \theta) = \begin{cases} \exp(i\theta) & |\vec{k}| \leq b/2 \\ 1 & |\vec{k}| > b/2, \end{cases} \quad (3.3)$$

here  $b$  is related with diameter of circular phase shift area  $a$  by  $b = \frac{2\pi}{\lambda} a$ , and  $\theta$  the phase difference induced by the phase plate. The output intensity distribution  $I_d = |U_2|^2$  is recorded by the photon detector.



(a)



(b)

Figure 3.2 (a) Input phase function for a single phase jump. (b) The transform function of the Zernike phase plate. Here  $a$  is the size of the phase shift area,  $\theta$  is the phase shift induced by the Zernike phase plate within the phase shift area.

While a measurement of a piston  $\phi$  in ZEUS is made on the assumption of a linear relation [15] between the sine function of this phase and its corresponding peak-to-peak (PtP) value of a localised intensity variation in the output and, as discussed in introduction, such a relation is an approximation, the degree of which depends on several factors such as the position and number of phase jumps, the aperture of the filter

and the size of the phase shift area of the filter. Here we examine this relation for ZEUS in the application in the segment mirror phasing and from this establish the measurement accuracy of the system. We begin with a single phase jump in the infinite one-dimensional space.

For a single phase jump in infinite one-dimensional space, the wave takes the form of:

$$U_1(x) = A_0 \exp[i\phi(x)], \quad (3.4)$$

where  $\phi(x)$  is the input phase function. The arbitrary single phase jump at  $x_1$  can be mathematically written as  $\phi(x) = \phi_1 + (\phi_2 - \phi_1)H(x - x_1)$ , the profile of which is shown in Fig.3.2(a). Here  $H(x)$  denotes the Heaviside function, with  $H(x) = 0$  for  $x < 0$  and  $H(x) = 1$  for  $x \geq 0$ . The transform function of the phase plate (Eq.3.3) introduced by the Zernike phase plate can be written as:

$$t(k) = [H(k + b/2) - H(k - b/2) + \frac{1}{e^{i\theta} - 1}](e^{i\theta} - 1). \quad (3.5)$$

As the diameter of the phase shift area  $b$  corresponds typically to  $\alpha = 1 - 2''$  on the sky [15], the relation between  $\alpha$  and  $b$  can be given as  $b \approx 2\pi L\alpha/\lambda$ , where  $L$  is the focal length. If we take the focal length of the first lens in Zernike filter system as 0.3m and the operating wavelength of 800nm, the value of  $b$  is around 600-1200. The intensity output of the ZEUS can be obtained as

$$I_d(x, \theta) = 2A_0^2 \sin \theta \sin(\phi_2 - \phi_1) \{ \text{sign}(x - x_1) - \text{Si}[b(x - x_1)/2]/\pi \} - \{ 4A_0^2 (1 - \cos \theta)(1 - \cos \phi) \{ 3 | \text{Si}[b(x - x_1)/2] | / \pi - 2\text{Si}^2[b(x - x_1)/2]/\pi^2 - 1 \} - A_0^2 \}. \quad (3.6)$$

where  $\text{sign}(x)$  is the sign function with  $\text{sign}(x) = 1$  for  $x \geq 0$  and  $\text{sign}(x) = -1$  for  $x < 0$ , the function  $\text{Si}$  is the Sinc integral:

$$\text{Si}(u) = \int_0^u \frac{\sin(t)}{t} dt. \quad (3.7)$$

Eq.3.6 comprises two parts. The anti-symmetrical part is used in ZEUS as information carrier for phasing [15], and the symmetric part is seen as a source of pollution to the signal.

### 3.3 Differential Zernike Filter and DZEUS

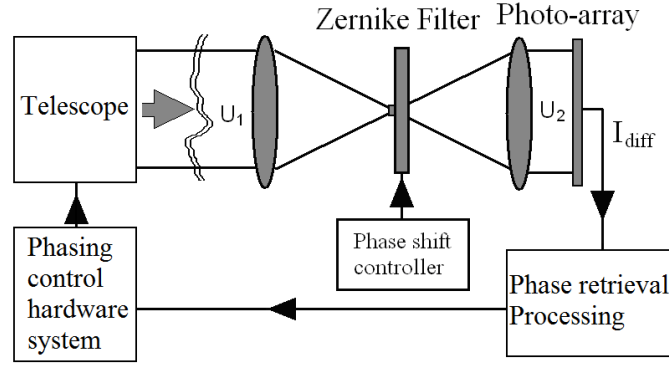


Figure 3.3 Schematic of the DZEUS phasing camera integrated in the Segmented Telescope.

DZEUS is realized by replacing the traditional Zernike filter with a differential Zernike filter in ZEUS. A schematic of differential Zernike filter in the DZEUS configuration is shown in Fig.3.3. The only difference in DZEUS system is that there is phase shift controller applied to the phase shift plate, which shift between  $\pm \theta$ . The differential Zernike filter was originally introduced by Vorontsov for wavefront sensing and correction [29]. The differential Zernike filter can be built by using a controllable phase shifting plate containing a single LC (liquid crystal) or MEMS (Microelectromechanical systems) actuator interfaced with the output photo array and image-subtraction system. The output intensity of differential Zernike filter is obtained by subtracting the two images recorded according to different phase shift in the phase plat, i.e.,

$$I_{diff}(\vec{r}) = [I_d(\vec{r}, \theta) - I_d(\vec{r}, -\theta)] / 2. \quad (3.8)$$

where  $I_d(x, \pm\theta)$  is the intensity output of the traditional Zernike filter for a phase shift of  $\pm \theta$ . Then the expression of the output intensity of differential Zernike filter can be derived by combining Eq.3.6 and 3.8:

$$I_{diff}(x) = A_0^2 \sin \theta \sin(\phi_2 - \phi_1) \{ \text{sign}(x - x_1) - \text{Si}[b(x - x_1)/2]/\pi \}, \quad (3.9)$$

We can see that the expression for  $I_{diff}$  is much simpler than that of  $I_d$ . This is because there are two parts in Eq. 3.6, the first part is anti-symmetrical and used in ZEUS as information carrier for phasing [15], and the second part is symmetrical which

is seen as a source of pollution to the signal. The subtraction used in DZEUS removes the symmetrical term, leading to a simplified result as shown in Eq. 3.9. An example of the output signal intensity profile of differential Zernike filter ( $I_{diff}$ ) and traditional Zernike filter ( $I_d$ ) are plotted in Fig. 3.4. The symmetrical term in  $I_d$  (dotted line), which is exactly the difference between  $I_d$  and  $I_{diff}$  is also shown in Fig. 3.4 (b) along with  $I_d$ .

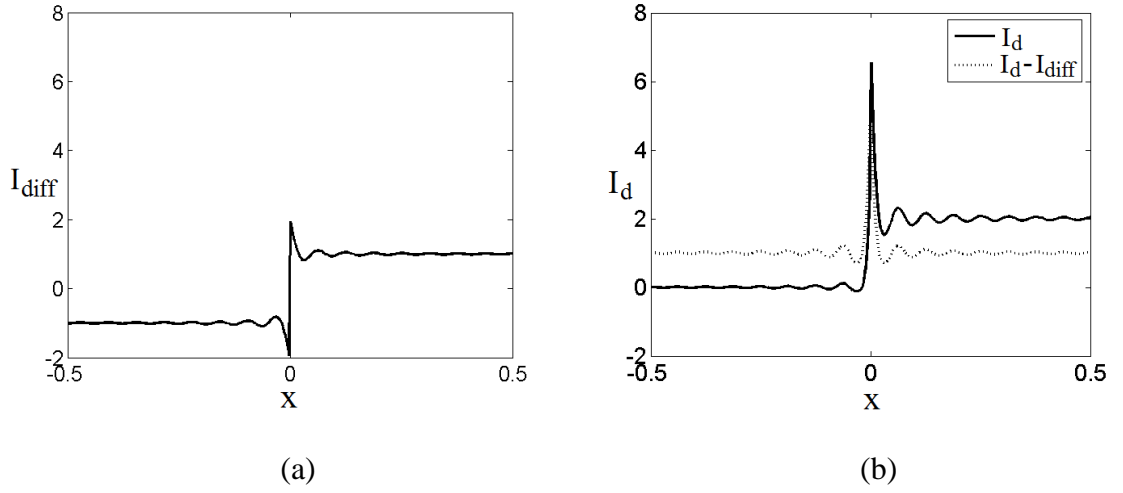


Figure 3.4 The output intensity of differential Zernike filter (a) and traditional Zernike filter (b) as a function of  $x$  with  $A_0 = 1$ ,  $\phi_2 - \phi_1 = \pi/2$ ,  $b = 200$ ,  $\theta = \pi/2$ , and  $x_1 = 0$ , obtained by analysis results with unlimited boundary.

The PtP amplitude in  $I_{diff}$ , which can be determined by searching for positive and negative peaks near the segment edge, is defined as:

$$PtP = \lim_{x \rightarrow x_1^+} I_{diff}(x) - \lim_{x \rightarrow x_1^-} I_{diff}(x) = 2A_0^2 \sin \theta \sin(\phi_2 - \phi_1). \quad (3.10)$$

We note that although the output profile of differential Zernike filter is different from that of Zernike filter, the PtP values for both are the same. However, this is true only for the current case of infinite aperture size. The value of piston error in the wavefront is given by:

$$p = \phi \frac{\lambda}{2\pi} = \sin^{-1}(PtP / A) \frac{\lambda}{2\pi} \quad (3.11)$$

This function can be used for the phasing errors (piston) retrieval.



### 3.4 Finite Aperture Size

In reality, the effects of finite sizes of both aperture and lenses should be taken into account. As higher frequency scattering (which is out of the lenses collection) is insignificant and can be ignored, we only need to consider the boundary effects caused by the limited size of the telescope aperture. In this case the input field amplitude is therefore no longer a constant but a function of position, which can be written as  $A(x) = A_0[H(x + \frac{d}{2}) - H(x - \frac{d}{2})]$ , where  $d$  ( $d > 0$ ) is the diameter of the segmented primary mirror. We consider only a single phase jump at  $x_1$ , which corresponds to the position of the inter-segment edge, so  $-\frac{d}{2} < x_1 < \frac{d}{2}$ . The output intensity of differential Zernike filter with a single phase jump in the input is:

$$\begin{aligned}
 I_{diff}(x) = & 2A_0^2 \sin \theta \sin(\phi_2 - \phi_1) \\
 & \{ \{ Si[b(x + d/2)/2]/\pi - Si[b(x - d/2)/2]/\pi \} H(x - x_1) \\
 & + \{ Si[b(x - x_1)/2]/\pi - Si[b(x + d/2)/2]/\pi \} H(x - d/2) \\
 & + \{ Si[b(x - d/2)/2]/\pi - Si[b(x - x_1)/2]/\pi \} H(x + d/2) \}.
 \end{aligned} \tag{3.12}$$

Eq. 3.12 is more complex than Eq. 3.9 as a result of the aperture. The output comprises two parts. The first part, with the function  $H(x - x_1)$ , is the carrier of the intersegment phase jump information. The second part, containing  $H(x - d/2)$  and  $H(x + d/2)$ , is the carrier of the boundary information. Eq. 3.12 can be reduced to Eq. 3.9 when the boundary size  $d$  goes to infinite. Moreover, the output intensity  $I_{diff}$  is no longer anti-symmetric in the presence of aperture. We plot  $I_{diff}$  against  $x$  in Fig. 4. For comparison, we also plot the result of traditional Zernike filter ( $I_d$ ) with same aperture to show the different characteristics of the two. A major difference is that the aperture profile appears in the intensity distribution of traditional Zernike filter, but not in that of differential Zernike filter. This is because the profiles are the same in outputs with  $\pm \pi/2$  phase shift and are cancelled by the differential algorithm in differential Zernike filter.

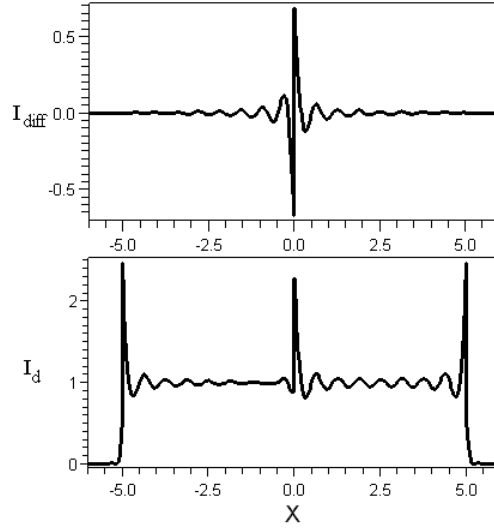


Figure 3.5 The output intensity of differential Zernike filter (top figure) and traditional Zernike filter (low figure ) as a function of  $x$  with  $A_0 = 1$ ,  $\phi_2 - \phi_1 = \pi/4$ ,  $b = 20$ ,  $c = \pi/2$ ,  $d = 10$ , and  $x_1 = 0$ .

The peak-to-peak (PtP) value of  $I_{diff}$  in the differential Zernike filter can be obtained as

$$\begin{aligned}
 PtP(x_1) &= \lim_{x \rightarrow x_1^+} I_{diff}(x) - \lim_{x \rightarrow x_1^-} I_{diff}(x) \\
 &= 2A_0^2 \sin \theta \sin(\phi_2 - \phi_1) \{ Si[b(x_1 + d/2)/2]/\pi - Si[b(x_1 - d/2)/2]/\pi \}.
 \end{aligned} \tag{3.13}$$

Here the PtP depends not only on the value of the input phase jump but also the size of the aperture  $d$ , the size of the phase shift area  $b$ , and the position of the phase jump  $x_1$ . This means that though the relation  $PtP \propto \sin(\phi_2 - \phi_1)$  still holds for the current case, the proportionality coefficient varies with the phase jump position for fixed value of  $b$  and  $d$ . The change of the coefficient is discussed in Fig. 3.6 for different value of phase shift area  $b$ . The PtP oscillates around  $2A_0^2 \sin \theta \sin(\phi_2 - \phi_1)$  (the value of which is 2 for the given parameter) on change of  $x_1$ . The oscillation amplitude decreases with the increase of  $b$  because of the dependence of the term  $Si[b(x_1 + d/2)/2]/\pi - Si[b(x_1 - d/2)/2]/\pi$  in Eq. 3.13 on  $b$ . Since  $b$  has a typical value of 600-1200 as discussed earlier, if  $x_1$  is away from  $\pm d/2$ , we have

$Si[b(x_1 + d/2)/2]/\pi - Si[b(x_1 - d/2)/2]/\pi \approx 1$ . Under this approximation, Eq. 3.13 can be written as:

$$PtP(x_1) \approx 2A_0^2 \sin \theta \sin(\phi_2 - \phi_1). \quad (3.14)$$

This is the same as the PtP expression obtained in the case of infinite aperture. The phase retrieval algorithm based on Eq. 3.11 can be used as a good approximation as long as  $b \gg 1$ .

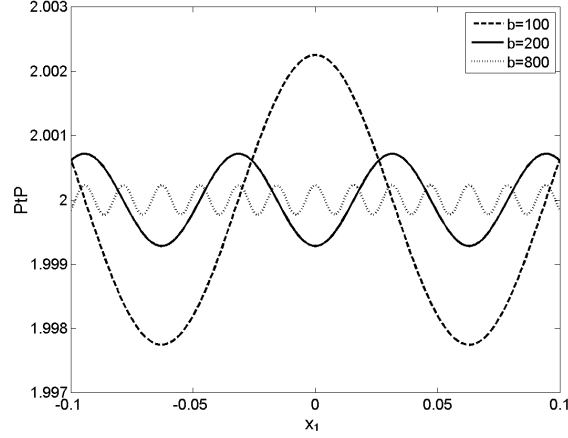


Figure 3.6 PtP value in the output signal of differential Zernike filter with single phase jump as a function of the position of the phase jump for different size of the phase mask  $b$  with  $A_0 = 1$ ,  $\theta = \pi/2$ ,  $\phi_2 - \phi_1 = \pi/2$ ,  $b = 100, 200, 400$ ,  $d = 20$ .

For comparison, we also give the analytic results for ZEUS in the presence of the same aperture. Since the derivative in this case is quite involved, we only write down the PtP value, which is

$$\begin{aligned} PtP(x_1) &= I_d(x \rightarrow x_1^+) - I_d(x \rightarrow x_1^-) \\ &= 2A_0^2 \sin \theta \sin(\phi_2 - \phi_1) \{ Si[b(x_1 + d/2)/2]/\pi - Si[b(x_1 - d/2)/2]/\pi \} \\ &\quad + 4A_0^2 (1 - \cos \theta) [1 - \cos(\phi_2 - \phi_1)] \{ Si[b(x_1 + d/2)/2]/\pi + Si[b(x_1 - d/2)/2]/\pi \}. \end{aligned} \quad (3.15)$$

while the first term in Eq. 3.15 is same as PtP value for differential Zernike filter, the second is additional and a function of  $\cos(\phi_2 - \phi_1)$ . As a result, the simple sine function relationship between PtP and phase jumps no longer hold for ZEUS in the case of finite aperture. We compare the performance of the two filters using an example in Fig. 3.7, where trace (a) gives the PtP values in DZEUS (dash) and ZEUS (solid) against phase jump positions. The fluctuation of the former is smaller than that of the latter due to the

presence of the second terms in Eq. 3.15, which is shown as the dash-dot curve in the same figure. In Fig. 3.7 (b), we plot the PtP values in DZEUS (dash) and ZEUS (solid) and their difference (dash dot) with respect to the value of  $\sin(\phi_2 - \phi_1)$ . While a linear relation holds very well for DZEUS, a deviation can be clearly seen for ZEUS, which can lead to increased measurement errors discussed later. From this analysis we conclude that in the case of a single phase jump DZEUS provide not only a simpler expression for the output intensity but also a simpler relation between PtP value and the input phase jump, which can be well approximated as  $\text{PtP} \propto \sin(\phi_2 - \phi_1)$  when  $b \gg 1$ .

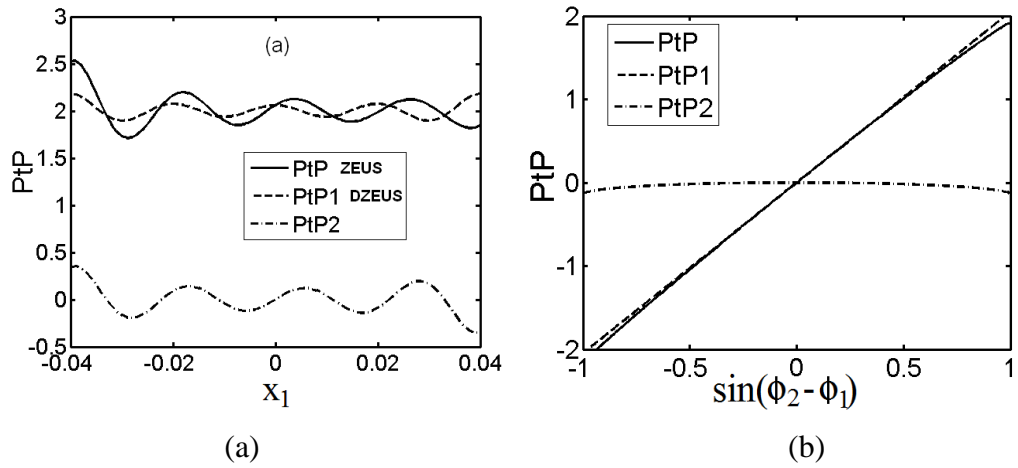


Figure 3.7 Value of the first term (PtP1, dash) and second term (PtP2, dash-dot) in Eq. 3.15 are plotted against  $x_1$  (a) and the value of  $\sin(\phi_2 - \phi_1)$  (b). The parameters used here are  $A_0 = 1$ ,  $\theta = \pi/2$ ,  $\phi_2 - \phi_1 = \pi/2$ ,  $b = 400$ ,  $d = 0.1$ .  $x_1 = 0.17$ .

### 3.5 Multiple Phase Jumps

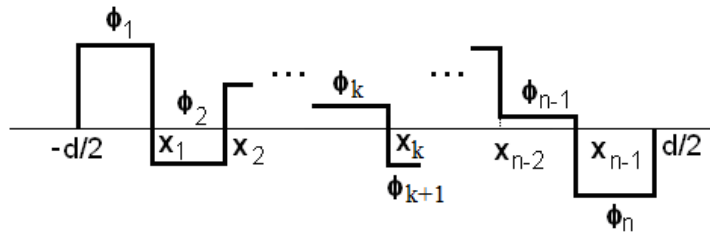


Figure 3.8 The input phase function with  $n$  segments and  $n-1$  phase jumps. Each segment has a random piston value.

The following analysis within this chapter is all conducted with finite primary aperture size if not mentioned otherwise. Based on the above argument, we extend out analysis to multiple phase jumps in a segmented mirror. In order to find the inter-segment effects between the nearby phase jumps, we consider the one dimensional model with a primary mirror consisting of  $n$  segments and thus having  $n-1$  inter-segment phase jumps as shown in Fig. 3.8. The PtP value in the output intensity for  $k$ -th ( $1 < k < n-1$ ) jump can be written as:

$$\begin{aligned}
PtP(x_k) = & \frac{2}{\pi} A_0^2 \sin \theta \{ [\sin(\phi_1 - \phi_k) + \sin(\phi_{k+1} - \phi_1)] \{ Si(\frac{kbd}{2n}) - Si[\frac{(k-1)bd}{2n}] \} \\
& + [\sin(\phi_2 - \phi_k) + \sin(\phi_{k+1} - \phi_2)] \{ Si[\frac{(k-1)bd}{2n}] - Si[\frac{(k-2)bd}{2n}] \} + \dots \\
& + [\sin(\phi_n - \phi_k) + \sin(\phi_{k+1} - \phi_n)] \{ Si[\frac{(k-n+1)bd}{2n}] - Si[\frac{(k-n)bd}{2n}] \} \}, \tag{3.16}
\end{aligned}$$

where  $\phi_i$  is the piston in the  $i$ -th segment,  $\phi(x) = \phi_i$  for  $x \in (x_{i-1}, x_i)$ ,  $-\frac{d}{2} < x_1 < x_2 \dots < x_{n-1} < \frac{d}{2}$ ,  $1 \leq k \leq n-1$ . Since the interval between the neighbouring segment edges are all equal,  $x_k = (k - \frac{n}{2}) \frac{d}{n}$ . We can see from Eq. 3.16 that the PtP value for the  $k$ -th phase jump is influenced not only by the  $k$ -th jump but also all other jumps in the segmented mirror. This makes the PtP value of multiple input phase jumps very complicated. The expression can however be significantly simplified in applications for a typical value  $b \gg 1$ . When  $b \gg 1$ , there is little difference between  $Si[\frac{(k-i+1)bd}{2n}]$  and  $Si[\frac{(k-i)bd}{2n}]$  in Eq. 3.16 when  $i \neq k$  or  $k+1$ , so all the terms in Eq. 3.16 except the  $k$ -th and  $k+1$ -th term can be approximated to zero. This simplifies Eq. 3.16 to

$$PtP(k) = \frac{4}{\pi} A_0^2 \sin \theta \sin(\phi_{k+1} - \phi_k) Si(\frac{bd}{2n}) \tag{3.17}$$

For  $b \gg 1$ ,  $Si(\frac{bd}{2n})$  can be approximated to  $\pi/2$ , thus

$$PtP(k) = 2A_0^2 \sin \theta \sin(\phi_{k+1} - \phi_k) \tag{3.18}$$

This is identical as Eq. 3.14, which we derived for the single phase jump condition. Therefore under the condition  $b \gg 1$ , Eq. 3.11 can also be used to retrieve the phase jump for multiple input phase jumps.

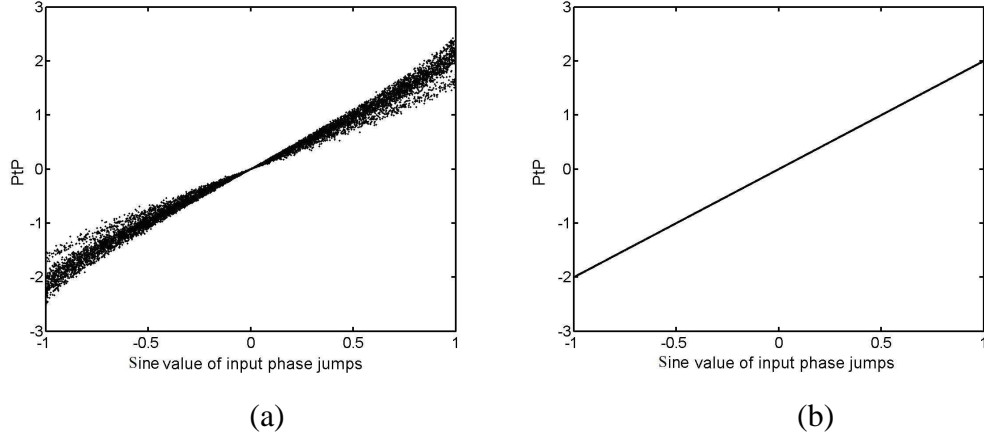


Figure 3.9 The PtP values in the output intensity distribution of differential Zernike filter versus the corresponding sine value of input phase jumps, with  $n = 20$ ,  $A_0 = 1$ ,  $c = \pi/2$ ,  $d = 20$ ,  $\phi_i$  are random numbers within  $[-\frac{\pi}{4}, \frac{\pi}{4}]$ , with 500 randomly different realization. (a)  $b = 2$ , (c)  $b = 400$ .

Let us now investigate the performance of DZEUS in a case study. We consider a segmented primary mirror of 20 segments, each of which has a diameter of 1m. The piston is randomly distributed within  $[-\frac{\pi}{4}, \frac{\pi}{4}]$  so the intersegment phase jumps are within  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ . The phase shift induced by the phase mask is  $\theta = \pi/2$ . The results are plotted in Fig. 3.9 (a) and (b) for 500 different (random) realizations for two phase mask sizes  $b = 2$  and 400, respectively. As seen, the linear relation holds well for  $b=400$  but not for  $b=2$ . In order to quantify the linearity of the relationship between the PtP value and sine function of the input phase jumps, we plot the standard error of estimate (SEE) against  $b$  in Fig. 3.10. SEE measures the standard deviation between the simulated PtP values and the values obtained by Eq. 3.14, and a smaller value of SEE implies a better linear relation. We note that the oscillations in the SEE curve is the result of fluctuations of Sinc integral function.

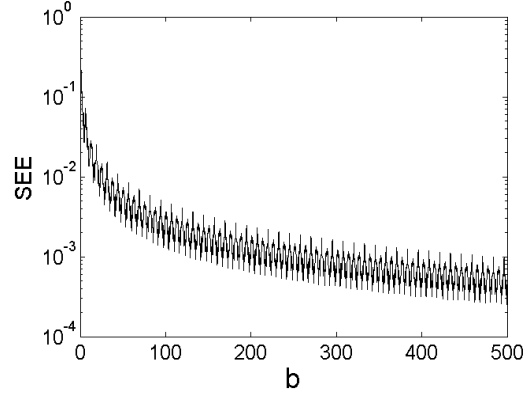


Figure 3.10 The standard error of estimate (SEE) of linear regression of the PtP values in the output intensity distribution of differential Zernike filter and the corresponding sine value of input phase jumps versus the parameter  $b$ . Other parameters used here are the same as in Fig. 3.9.

### 3.6 Double Phase Jumps

The expression for the output of traditional Zernike filter is too complicated to do any further analysis with arbitrary number of phase jumps. In order to make the multiple phase jumps output PtP value for differential Zernike filter comparable with that of traditional Zernike filter, we consider a segmented primary mirror with 3 segments. The input phase function is shown in Fig. 3.11. Here we keep  $\phi_1 = \phi_3 = -\phi_0/2$ , and  $\phi_2 = \phi_0/2$  ( $\phi_0 > 0$ ) for simplification, so we derive from Eq. 3.16:

$$PtP(x_1) = \frac{2}{\pi} A_0^2 \sin \theta \sin \phi_0 \{ Si[b(x_1 + d/2)/2] - Si[b(x_1 - d/2)/2] \}, \quad (3.19)$$

$$PtP(x_2) = \frac{2}{\pi} A_0^2 \sin \theta \sin \phi_0 \{ Si[b(x_2 - d/2)/2] - Si[b(x_2 + d/2)/2] \}. \quad (3.20)$$

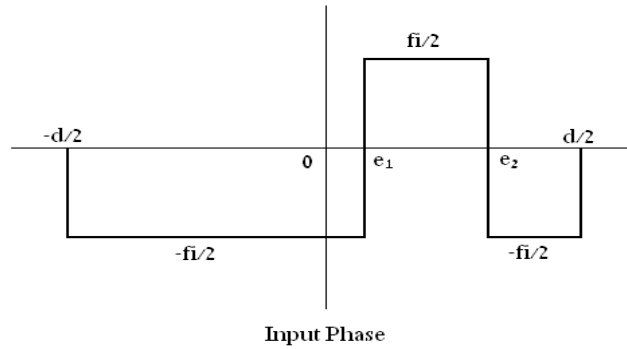


Figure 3.11 The input phase function with 3 segments and 2 phase jumps.

The expression of PtP value for differential Zernike filter output with two phase jumps is identical to that with single phase jump. However, the PtP values in the output intensity of traditional Zernike filter are different from that with single phase jump:

$$\begin{aligned}
PtP(x_1) = & \frac{2}{\pi} A_0^2 \sin \theta \sin \phi_0 \{ Si[b(x_1 + d/2)/2] - Si[b(x_1 - d/2)/2] \} \\
& + \frac{4}{\pi} A_0^2 (1 - \cos \theta)(1 - \cos \phi_0) \{ Si[b(x_1 + d/2)/2] - Si[b(x_1 - d/2)/2] \\
& + 2Si[b(x_1 - x_2)/2] \}, \tag{3.21}
\end{aligned}$$

$$\begin{aligned}
PtP(x_2) = & \frac{2}{\pi} A_0^2 \sin \theta \sin \phi_0 \{ Si[b(x_2 - d/2)/2] - Si[b(x_2 + d/2)/2] \} \\
& + \frac{4}{\pi} A_0^2 (1 - \cos \theta)(1 - \cos \phi_0) \{ Si[b(x_2 - d/2)/2] - Si[b(x_2 + d/2)/2] \\
& + 2Si[b(x_2 - x_1)/2] \}. \tag{3.22}
\end{aligned}$$

Again, the PtP value in the traditional Zernike filter output intensity is consists of two parts as well as single phase jump case. Besides, there is a cross term, which is a function of the distance of the two phase jumps' positions respectively in Eq. 3.21 and 3.22. That means the PtP value according to one phase jump rely on the distance between the neighbouring phase jumps. In other words, the existence of more than one phase jump can result a crosstalk between each PtP value in traditional Zernike filter output intensity, but not in that of differential Zernike filter. This crosstalk will further deteriorate the assumed simple relation used in the phase retrieval in ZEUS.

### 3.7 Multiple Step Corrections

From the above analysis, Eq. 3.11 can be used as an approximation to retrieve multiple segment phase jumps based on the measurement of the PtP value. The correction is nevertheless not precise and errors remain. Here we propose a multiple step correction approach by which the measurement and adjustment process is repeated several times until required accuracy is met. A necessary condition for such a process is the convergence of the system, which corresponds to the inequality

$$| \phi_{k+1} - \phi_k - \sin^{-1} \left[ \frac{PtP(x_k)}{2A_0^2 \sin \theta} \right] | < | \phi_{k+1} - \phi_k | \tag{3.23}$$



which is derived from the phase retrieval algorithm (Eq. 3.11). Since the exact expression for  $PtP(x_k)$  (Eq. 3.16) is complicated, it is hard to verify the inequality in general case. It can be easily done under the small piston variation approximation, i.e.,  $\sin(\phi_j - \phi_k) \approx \phi_j - \phi_k$ . Under this approximation, we prove that this system is convergence for small signal. In this case Eq. 3.16 is simplified as

$$PtP(x_k) = \frac{2}{\pi} A_0^2 \sin \theta (\phi_{k+1} - \phi_k) \left\{ Si\left(\frac{kbd}{2n}\right) - Si\left[\frac{(k-n)bd}{2n}\right] \right\} \quad (3.24)$$

By combine Eq. 3.23 and 3.24, we have

$$0 < Si\left(\frac{kbd}{2n}\right) + Si\left[\frac{(n-k)bd}{2n}\right] < 2\pi. \quad (3.25)$$

As  $0 < Si(x) < \pi$ , for any  $x > 0$ , the above inequality is valid unconditionally. The performance of such a multiple step correction procedure is shown in Fig. 3.12. The phasing error is indeed convergent to zero on increasing the iteration number. By comparison, we also provide the corresponding results for ZEUS, for which the error reaches a plateau after a few steps of correction. The latter is due to the effect of pupil profile under a finite resolution.

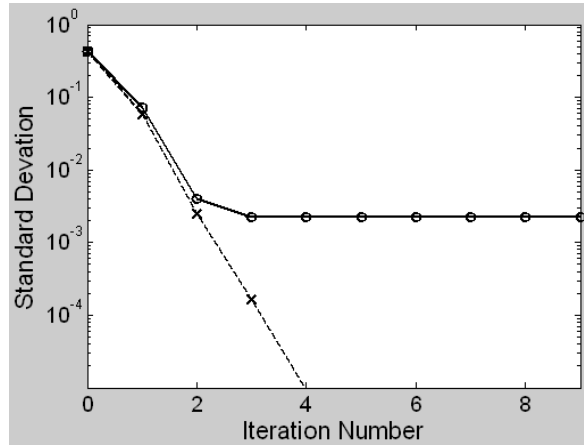


Figure 3.12 The standard deviation of the phasing error (radius) of the segmented mirror versus iteration numbers. The initial phase jumps are uniformly distributed within  $[-\pi/2, \pi/2]$ . The dash line with cross is for DZEUS and the solid line with circle for ZEUS. The resolution in the simulation is 1024 for 16 segments. Other parameters are  $A_0 = 1$ ,  $\theta = \pi/2$ ,  $b = 600$ ,  $d = 16$ .

### 3.8 Two-Dimensional Simulations

We now extend the one-dimensional analysis to two-dimensional case. Since the analysis is complicated, we deal with it through numerical simulations. We want to answer two questions here: whether the pseudo-linear relation obtained in DZEUS in one dimension case still holds for phasing of a segmented telescope (two-dimension), and whether the multiple step correction approach can be adopted to improve the correction accuracy until a desired accuracy?

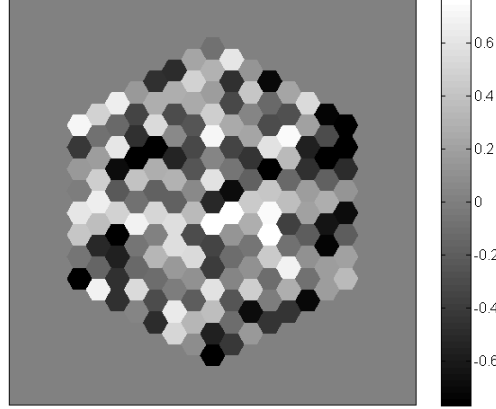


Figure 3.13 A snapshot of the segmented primary mirror used in simulation. The piston of each segment is normal randomly distributed. The grey scale represents the piston value of each segment, which is ranged from  $-\pi/4$  (darkest) to  $\pi/4$  (brightest).

The segmented mirror in our simulation is a 169-hexagon formation, each of which has a random piston. A snapshot of the segmented primary mirror is shown in Fig. 3.13. We consider each mirror segment of 1 m in diameter so the primary mirror in this simulation is 15 m in diameter. The input pistons of the primary mirror are randomly distributed within  $[-\pi/4, \pi/4]$ , so the intersegment phase jumps are distributed within  $[-\pi/2, \pi/2]$ , which results in a standard deviation of the input phase jump of 0.462 rad. For piston errors beyond this range, the ambiguity due to phase wrapping can be solved by the two wavelengths interferometry algorithm [26]. In our simulation we use  $1024 \times 1024$  pixels for the 169-hexagon formation. The resolution is the same as that for a CCD camera of the resolution  $4096 \times 4096$  for the European Extremely Large Telescope with a primary mirror of 42m. We set the size of the image pupil 0.1m, so one pixel in the image pupil corresponds to 0.2mm. We plot in Fig. 3.14 (a) the PtP values in the output intensity in DZEUS against the sine function of the corresponding

input phase jumps. The curve shows a good linear relation between the two. For comparison, we also plot the standard error of estimate (SEE) between the input phase jumps and the retrieved values for both DZEUS and ZEUS. As shown in Fig. 3.14 (b), the SEE values for DZEUS is consistently smaller than that of ZEUS on varying  $b$ , implying a better linear relation for the former for the usual operating region of  $b$  up to 1200. Beyond this region, the two curves converge because the relation between the input and output become identical for very larger value of  $b$ . For a typical size of the phase shift area  $b = 817$ , which corresponds to 13 pixel in our simulation, we have  $SEE = 0.0163$  for DZEUS, compared to 0.0338 ZEUS. The remaining errors after one step correction are 8.4% and 12.5% of the initial value for DZEUS and ZEUS respectively.

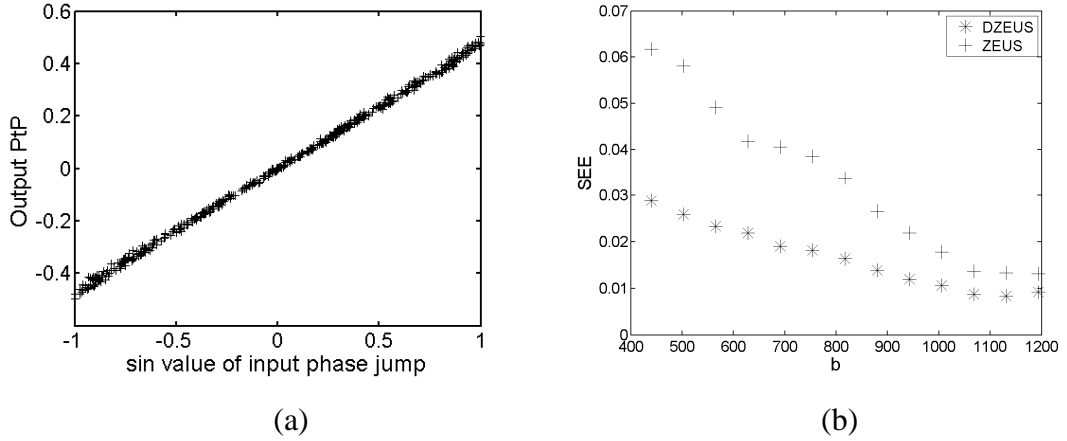


Figure 3.14 Two-dimensional simulation results with a primary mirror of 169 segments. The piston randomly distributed within  $[-\pi/4, \pi/4]$ . (a) The output PtP value versus the sine value of the input phase jumps. (b) The standard error of the estimate (SEE) versus the phase shift area size ( $b$ ).

Using multiple step correction procedure, the phasing error retrieved based on the above phase retrieval algorithm is used to correct intersegment errors through the actuator control system, and the process is repeated for several times to achieve high accuracy. The retrieved intersegment errors are transformed into pistons values by singular value decomposition<sup>1</sup>. In our simulation, we assume there is no noise and error in the hardware system. That means the retrieved phasing error is perfectly removed

<sup>1</sup> See Appendix A

from the segmented mirrors, and the remaining phasing errors are results of the phase retrieval algorithm. The two-dimensional simulations are conducted with respect to DZEUS and ZEUS for comparison. In Fig. 3.15, we plot the standard deviations of phasing errors against the iteration number for both DZEUS and ZEUS. As shown in Fig. 3.15, the phasing error reduces exponentially to zero on increasing iteration number for DZEUS but reaches to a plateau after 3 iterations for ZEUS. The behaviours of the two systems are therefore the same as them in one dimension as discussed earlier.

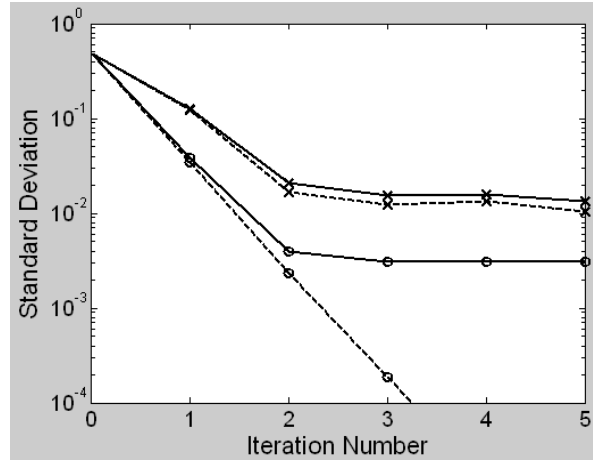


Figure 3.15 The standard deviation of the phasing error versus iteration numbers. The dash (solid) line with circle is for DZEUS in the absence (presence) of atmospheric turbulence whereas the dash (solid) line with cross is for ZEUS in the absence (presence) of atmospheric turbulence. Atmospheric turbulence is numerically generated at each exposure according to Komogorov model, the amplitude of which used in the plot is  $2\pi$ .

Finally we consider the effect of atmospheric turbulence on the measurement precision. When the turbulence is considered, a finite error remains for both algorithms, the value of which is lower for DZEUS compared to ZEUS as shown in Fig. 3.15. We note that the difference of the errors between DZEUS and ZEUS is kept almost a constant on the variation of the amplitudes of atmospheric turbulence. For a practical system with the central operating wavelength of 800nm, the best phasing accuracy is 1.4nm for DZEUS and 1.6nm for ZEUS. Simulation results show that the both algorithms are robust with respect to atmospheric turbulence. We note that reported

phasing error for ZEUS is higher than the one indicated in our simulation. [15] The discrepancy mainly results from errors in system hardware which has not been considered here.

### 3.9 Conclusions

In conclusion, we have demonstrated a new segmented mirror phasing sensor system, DZEUS, by theoretical and numerical study. We show that the DZEUS gives rise to a simpler relation between the input wave front and the output intensity compared to ZEUS and therefore provides better accuracy for the phase retrieval. The phasing accuracy can be further enhanced by adopting a multiple step correction approach, by which the phasing error in the ideal situation can be removed completely. In practice, phase switching in DZEUS can be realized by replacing the fixed mask array in ZEUS with an electronically controllable spatial light modulator (SLM). DZEUS inherits the advantages of ZEUS phasing technique by resemblance to the Mach-Zehnder interferometer, which avoids the delicacy of high-precision control of the OPD of interferometer. As for ZEUS, DZEUS provides a capture range of  $\pm\lambda/4$  in single correction and  $\pm\lambda/2$  in multiple step corrections and is robust to atmosphere turbulence.

## **Chapter 4**

### **Differential Zernike Feedback Phase Retrieval**

#### **4.1 Introduction**

In the last chapter, differential Zernike filter is used as a wavefront sensor to retrieve the segmented mirror phasing errors. The phasing errors are retrieved at each exposure, and the measured phasing errors are used to control the actuator. In this chapter, we present my original study on an adaptive dynamical feedback system, which is realized by combining the capabilities of a differential Zernike sensor with a feedback for phasing of segmented mirror in the presence of strong atmospheric turbulence. The dynamics of the feedback signal in the sensor system has the same form as that of a Gradient-Flow control [30], which results an optimized wavefront conjugation. A Gradient-Flow control (optimization) represents a method based on optimizing the system (signal) state by gradient metrics. The gradient-flow optimization method is widely used for digital image processing applications [39-41]. The gradient metrics is calculated analytically based on knowledge of the system's mathematical model and performance metric. The mechanism underlying the differential Zernike filter feedback technique for phasing of segmented mirror in the presence of strong atmospheric turbulence lies in the dynamical nature of the system. It responds to static misalignment phase errors differently from dynamic atmospheric turbulence. During the iterating process of the system, only the signals related to the static phase errors are accumulated in the feedback loop whereas those with the turbulent phases are essentially averaged

out, because the latter is uncorrelated from frame to frame. When the iteration has converged, the signal in the feedback loop is the conjugate of the misalignment phase errors, and so can be used to control the actuators of the segmented mirrors.

## 4.2 Theoretical Model

### 4.2.1 Zernike filter with Unknown Input Wavefront

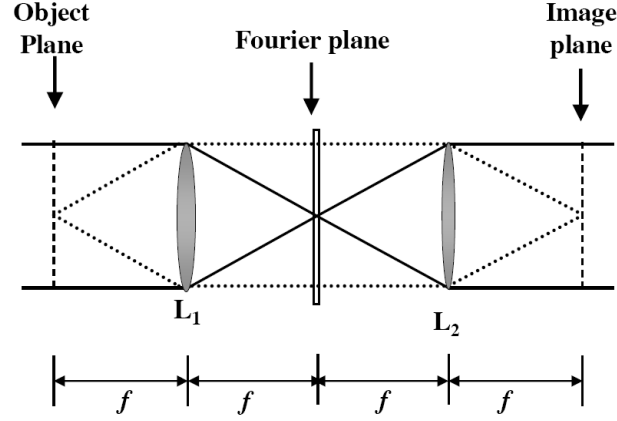


Figure 4.1 Schematic of Zernike filter  $4f$  optical system. Two identical lenses  $L_1$  and  $L_2$  are placed by twice of their focal length  $f$ .

In this chapter we use a different mathematical approach for Zernike filter by considering an unknown input wavefront. A schematic of a conventional wave-front sensor based on the Zernike phase-contrast technique (Zernike filter) is shown in Fig. 4.1. It consists of two lenses with a phase changing plate (Zernike phase plate) placed in the lenses' common focal plane. The phase plate has a small circular region (a dot) in the middle that introduces a phase shift  $\theta$  near  $\pi/2$  rad into the focused wave. The radius of the dot,  $a_F$ , is typically chosen to equal the diffraction limited radius  $a_F^{dif}$  of a focused, undistorted input wave. The treatment of the Zernike filter that we follow is presented in Ref. 29 and 32. To introduce the notation and normalizations, we offer the basic derivations that lead to the classical expression for the Zernike filter output intensity distribution. Zernike filter can be described by using a complex transfer function  $t(\vec{k})$  for the focal plane filter:

$$\begin{aligned}
t(\vec{k}) &= \gamma \exp(i\theta) \quad |\vec{k}| \leq k_0, \\
t(\vec{k}) &= 1 \quad |\vec{k}| > k_0,
\end{aligned} \tag{4.1}$$

The wave vector  $\vec{k}$  is associated with focal plane radial vector  $\vec{r}_F$  through  $\vec{k} = \vec{r}_F / (\lambda L)$ , where  $L$  is the lens focal length,  $\lambda$  is the central wavelength of the input light, and  $b = a_F / (\lambda L)$  is the cut-off frequency that corresponds to dot size  $a_F$ . For the sake of convenience, consider the following variable normalization: The radial vectors  $\vec{r}$  in the sensor input/output plane and  $\vec{r}_F$  in the focal plane are normalized by the lens aperture radius  $a$ , the wave vector  $\vec{k}$  by  $a^{-1}$ , and the lens focal length by the diffraction parameter  $ka^2$  (where  $k = 2\pi / \lambda$  is the wave number). Correspondingly, in the normalized variables,  $\vec{k} = \vec{r}_F / (2\pi L)$  and  $b = a_F / (2\pi L)$  (where the dot size  $a_F$  is also normalized by  $a$ ).

From Eq. 4.1, when  $\gamma = 1$  and  $\theta = \pi / 2$ , we have a Zernike filter model. Consider a simplified model corresponding to a focal plane filter affecting only the zero-order spectral component. In this case we have  $T(0) = \gamma \exp(i\theta)$ , and  $T(\vec{k}) = 1$  for  $\vec{k} \neq 0$ . Assume an input wave  $A_{in}(\vec{k}) = A_0(\vec{r}) \exp[i\varphi(\vec{r})]$  enters a wave-front sensor, where  $I_0(\vec{r}) = A_0^2(\vec{r})$  and  $\varphi(\vec{r})$  are the input wave intensity and phase spatial distributions. The sensor's front lens performs a Fourier transform of the input wave. Within the accuracy of a phase factor,  $A(\vec{k}) = (2\pi L)^{-1} F[A_{in}(\vec{r})]$ , where  $F[]$  is the Fourier transform operator and  $A(\vec{k})$  is the spatial spectral amplitude of the input field (i.e., the field complex amplitude in the focal plane). In normalized variables, the field intensity in the focal plane can be expressed as a function of spatial frequency:  $I_F(\vec{k}) = (2\pi L)^{-2} |A(\vec{k})|^2$ . The influence of the focal plane filter can be accounted for by multiplying  $A(\vec{k})$  by the transfer function  $T(\vec{k})$ :

$$A_{out}(\vec{k}) = A(\vec{k})[1 - \delta(\vec{k})] + \gamma \exp(i\theta) A(\vec{k}) \delta(\vec{k}) \tag{4.2}$$



where  $A_{out}(\vec{k})$  is the focal plane wave complex amplitude after the wave passes through the spatial filter and  $\delta(\vec{k})$  is a delta function. The wavefront sensor output field can be obtained by taking the inverse Fourier transform of Eq. 4.2:

$$A_{out}(\vec{r}) = A_{in}(\vec{r}) - [1 - \gamma \exp(i\theta)] \bar{A}, \quad (4.3)$$

$$\bar{A} = \int A_{in}(\vec{r}) d^2\vec{r}, \quad (4.4)$$

where  $\bar{A}$  is the spatially averaged input field complex amplitude. For the sake of simplicity the 180° rotation of the field performed by the wave-front sensor lens system is ignored.

Represent  $\varphi(\vec{r})$  as a sum of mean phase  $\bar{\varphi}$  and spatially modulated deviation  $\tilde{\varphi}(\vec{r})$  :  $\varphi(\vec{r}) = \bar{\varphi} + \tilde{\varphi}(\vec{r})$ . In this case  $\bar{A} = \exp(i\bar{\varphi}) \bar{A}_0$  , where  $\bar{A}_0 = \int A_0(\vec{r}) \exp[i\tilde{\varphi}(\vec{r})] d^2\vec{r}$ . The value of  $|\bar{A}_0|^2$  is proportional to the field intensity at the centre of the lens focal plane,  $I_F(\vec{q}=0) = (2\pi L)^{-2} |\bar{A}_0|^2$  (intensity of the zero-order spectral component). The normalized value of  $I_F(0)$  is known as the Strehl ratio,  $St = I_F(0)/I_F^0$ , where  $I_F^0$  is the intensity of the zero-order spectral component in the absence of phase aberrations. When there is no phase aberration,  $I_F(0) = I_F^0$ , so  $St = 1$ . When there is any phase aberration,  $I_F(0) < I_F^0$  ( $St < 1$ ). With the notation introduced here, Eq. 4.3 can be written as

$$A_{out}(\vec{r}) = A_{in}(\vec{r}) - [1 - \gamma \exp(i\theta)] \bar{A}_0 \exp(i\bar{\varphi}). \quad (4.5)$$

As we can see from Eq. 4.5, the output field is a superposition of the input and spatially uniform reference wave components. Represent the complex value  $\bar{A}_0$  in the following form:  $\bar{A}_0 \equiv |\bar{A}_0| \exp(i\Delta) = (2\pi L) I_F^{1/2}(0) \exp(i\Delta)$ , where  $I_F^0$  and  $\Delta$  are the intensity and the phase, respectively, of the zero-order spectral component. The intensity distribution in the Zernike filter output plane for an unknown input wavefront is given by

$$\begin{aligned}
I_d(\vec{r}) = & I_0(\vec{r}) + (2\pi L)^2 I_F(0)(1 + \gamma^2 - 2\gamma \cos \theta) \\
& - 4\pi L I_0^{1/2}(\vec{r}) I_F^{1/2}(0) \{ \cos[\tilde{\varphi}(\vec{r}) - \Delta] \} \\
& - \gamma \cos[\tilde{\varphi}(\vec{r}) - \Delta - \theta].
\end{aligned} \tag{4.6}$$

#### 4.2.2 Differential Zernike Filter

A phase spatial light modulator (SLM) can be used to implement the differential Zernike filter [29]. The differential Zernike filter that is realized by using a phase shift switching between  $\pm \pi / 2$ , the output of which is given as

$$\begin{aligned}
I_{diff}(\vec{r}, t) = & \frac{1}{2} [I_d^{(+)}(\vec{r}, t) - I_d^{(-)}(\vec{r}, t)] \\
= & 2A_0(\vec{r}, t) |\bar{A}_0(t)| \sin[v(\vec{r}, t) - \Delta],
\end{aligned} \tag{4.7}$$

where  $I_d^{(+)}(\vec{r}, t)$  and  $I_d^{(-)}(\vec{r}, t)$  are the images recorded corresponding to the  $\pm \pi / 2$  phase shift in the phase plate,  $\bar{A}_0(t) = |\bar{A}_0(t)| \exp(i\Delta) = \int A_0(\vec{r}, t) \exp\{i[v(\vec{r}, t)]\} d^2\vec{r}$ , here  $A_0(\vec{r}, t)$  and  $v(\vec{r}, t)$  are the amplitude and phase of the input field,  $\Delta$  is a phase shift,  $\vec{r}$  is the spatial radial vector in the plane transverse to the system optical axis and the integration is over the aperture area.

### 4.3 Wavefront Control Based on Gradient-Flow Optimization

The differential Zernike filter can be combined with an adaptive feedback system, leading to applications such as wavefront compensation without the requirement of wavefront reconstruction. The differential Zernike filter offers a means for implementing a direct-control adaptive optical system by use of the gradient-flow optimization, which results in dramatic improvement in adaptation process convergent speed [30].

#### 4.3.1 Feedback-Controller Synthesis

First we consider a direct-control adaptive optics system [30] shown in Fig. 4.2. This system consists of the following adaptive optics components: wave-front corrector, wave-front sensor, and feedback controller. All of the adaptive system components are assumed to have high spatial resolution, and thus a continuously distributed

approximation of the adaptive system model can be used. The wave-front corrector (can be a spatial light modulator) introduces a phase modulation  $u(\vec{r}, t)$  into the distorted input wave  $A_{in}(\vec{r}, t) = A_0(\vec{r}) \exp[i\phi(\vec{r}, t)]$ . The corrected wave  $A_{cor}(\vec{r}, t) = A_0(\vec{r}) \exp\{i[\phi(\vec{r}, t) + u(\vec{r}, t)]\}$  is used as the wave-front sensor input. The wave-front sensor is interfaced with the feedback controller, which operates directly by using the sensor's output intensity  $I_{out}(\vec{r}, t)$ .

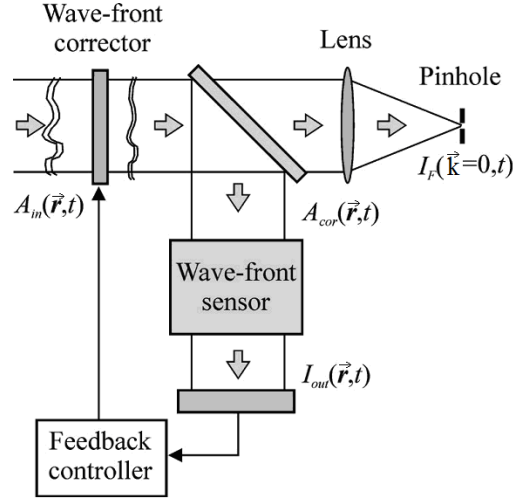


Figure 4.2 Schematic of a direct-control adaptive optics system.

The dependence of the correction function  $u$  on the wave-front sensor output  $I_{out}$  defines the control algorithm of feedback controller. For a continuous-time controller this algorithm can be represented as a time-dependent controlling phase-evolution process:

$$\frac{\partial u(\vec{r}, t)}{\partial t} = G(u, I_{out}), \quad (4.8)$$

in which  $G$  is an operator describing the feedback controller.

Synthesis of the wave-front controller  $G$  can be based on different principles. In the diffractive-feedback adaptive system, both the wave-front sensor and the controller  $G$  are selected on the basis of an analysis of the nonlinear spatiotemporal dynamics of equation 4.8 [34, 35]. The major requirement for these dynamics, or equivalently for the feedback controller design, is the existence of stationary state solutions that correspond to phase distortion suppression.

Another approach to wave-front controller synthesis is based on the gradient

optimization technique [36–38]. In this case the control rule equation 4.8 describes a continuous-time gradient-descent optimization of a system performance metric  $J$  :

$$\frac{\partial u(\vec{r}, t)}{\partial t} = \eta J'(\vec{r}, t), \quad (4.9)$$

where  $J'(\vec{r}, t)$  is a first variation (gradient) of the cost function and  $\eta$  is a constant positive for cost functional maximization and negative otherwise.

For practical implementation of the gradient-flow technique in adaptive optics, the gradient  $J'(\vec{r}, t)$  should be dependent only on available information: Here, the wave-front sensor output intensity  $I_{out}(\vec{r}, t)$  and the controlling phase  $u(\vec{r}, t)$ . This requires gradient representation in the following form:  $J'(\vec{r}, t) = J'[u, I_{out}]$ . As is shown in the following sections, the gradient-flow direct-control technique can indeed be used for adaptive wave-front distortion correction.

#### 4.3.2 System Performance Metric and Gradient-Flow Dynamics

For a number of adaptive optics applications (e.g., imaging of point-source objects, laser communication), a natural measure of system performance in correcting the distorted wave front is the Strehl ratio. Maximization of the Strehl ratio by the gradient-descent technique may result in two undesirable phenomena: drift of the aperture-averaged phase  $\bar{u}(t)$  toward the edge of the wave-front corrector's operational range, and phase discontinuities, both of which may occur during the adaptation process. To prevent aperture-averaged phase drift and to smooth the controlling phase (i.e., to suppress discontinuities and noise) the system performance metric  $J$  may include (besides the Strehl ratio) additional penalty terms:

$$J[u] = St - \alpha_1 [\bar{u}(t) - u_0]^2 - \alpha_2 \int |\nabla u(\vec{r}, t)|^2 d^2 r. \quad (4.10)$$

where  $\bar{u}(t) = S^{-1} \int u(\vec{r}, t) d^2 r$  is the phase averaged over the aperture area  $S$ ,  $u_0$  is a desirable value for  $\bar{u}(t)$ , and  $\alpha_1$  and  $\alpha_2$  are weight coefficients determining penalty term contributions. For now, ignore in Eq. 4.10 the time dependence of both phase aberrations and the controlling phase by assuming that phase aberrations are stationary. The complex amplitude of the input field (after it passes through the wave-front

corrector) can then be represented in the form  $A_{cor}(\vec{r}) = A_0(\vec{r}) \exp\{i[u(\vec{r}) + \tilde{\varphi}(\vec{r})]\}$ .

Then equation 4.10 can be changed as

$$J[u] = \int A_0(\vec{r}) \exp\{i[u(\vec{r}) + \tilde{\varphi}(\vec{r})]\} d^2r \left| \right|^2 - \beta_1 [u_0 - S^{-1} \int u(\vec{r}) d^2r]^2 - \beta_2 \int |\nabla u(\vec{r})|^2 d^2r \quad (4.11)$$

where  $\tilde{\varphi}(\vec{r}) = \varphi(\vec{r}) - \bar{\varphi}$  is the spatially modulated component of the wave-front aberration and  $\beta_1$  and  $\beta_2$  are new weight coefficients. The first term in Eq. 4.11 is proportional to the intensity of the input field's zero-order spectral component  $I_F(\vec{k} = 0)$ . Note that expressions for the weighting coefficients in Eq. 4.11 are irrelevant for the analysis below and for this reason are not defined. Consider the variation  $\delta J$  of the cost functional resulting from the small perturbation  $\delta u$  of the controlling phase:

$$\delta J = J[u + \delta u] - J[u] = \int J'(\vec{r}) \delta u(\vec{r}) d^2r + o(\delta u), \quad (4.12)$$

where the term  $o(\delta u)$  describes second- and higher-order terms with respect to the phase variation  $\delta u$ . Using Eq. 4.12 for the cost functional gradient we obtain

$$J' = -2 |\bar{A}_0| A_0(\vec{r}) \sin[u(\vec{r}) + \tilde{\varphi}(\vec{r}) - \Delta] - 2\beta_1 (\bar{u} - u_0) + 2\beta_2 \nabla^2 u(\vec{r}). \quad (4.13)$$

Here

$$\bar{A}_0 \equiv |\bar{A}_0| \exp(i\Delta) = \int A_0(\vec{r}) \exp\{i[u(\vec{r}) + \tilde{\varphi}(\vec{r})]\} d^2r. \quad (4.14)$$

Note that the value  $|\bar{A}_0|^2$  in Eq. 4.13 is proportional to the Strehl ratio.

Embed the control function  $u(\vec{r})$  in a family of time-dependent functions  $u(\vec{r}, t)$ , and consider the time-dependent evolution of  $J$  in the direction of the cost functional gradient. Thus the gradient-flow dynamics describes by Eq. 4.9 leads to the following nonlinear diffusion equation describing the controlling phase update:

$$\frac{\partial u(\vec{r}, t)}{\partial t} = d \nabla^2 u(\vec{r}, t) - \gamma |\bar{A}_0(t)| A_0(\vec{r}) \sin[u(\vec{r}, t) + \tilde{\varphi}(\vec{r}, t) - \Delta] - \mu [\bar{u}(t) - u_0], \quad (4.15)$$

where  $d$ ,  $\gamma$ , and  $\mu$  are coefficients dependent on the parameters  $\alpha_1$ ,  $\alpha_2$ , and  $\eta$  introduced in Eq. 4.9 and 4.10. We note that this differential equation has infinite

numbers of resolutions. This ambiguity can be solved by running this system with two different wavelengths. The details will be provided in this chapter.

### 4.3.3 Gradient-Flow Dynamics and Differential Zernike filter Synthesis

Compare the sinusoidal term in the formula for the gradient (Eq. 4.15) with the expression for the output intensity for the differential Zernike wave-front sensor (Eq. 4.7). The key observation from this comparison is that the gradient (Eq. 4.13) can be represented in a form dependent only on the differential Zernike filter output intensity  $I_{diff}(\vec{r}, t)$  and controlling phase  $u(\vec{r}, t)$ . Then, Eq. 4.15 for the feedback controller based on the differential Zernike filter can be written as

$$\frac{\partial u(\vec{r}, t)}{\partial t} = d\nabla^2 u(\vec{r}, t) - KI_{diff}(\vec{r}, t) - \mu[\bar{u}(t) - u_0], \quad (4.16)$$

where the coefficient  $K$  is proportional to  $\gamma$  in Eq. 4.15. The most important conclusion from this analysis is the following: The feedback controller Eq. 4.16 is an implementation of continuous-time gradient-flow dynamics, leading to a maximization of the cost functional (Eq. 4.11). It can be shown that during the adaptation process, the time derivative of  $J$  is always positive ( $dJ/dt > 0$ ); that is, the feedback controller (Eq. 4.16) provides for a monotonic increase in Strehl ratio, or in another word, decrease in the wavefront aberration.

## 4.4 Differential Zernike Feedback Phasing Sensor

### 4.4.1 System Model of Differential Zernike Feedback Sensor

Fig. 4.3 shows the schematic of the entire telescope system including the differential Zernike filter feedback wavefront sensor. Our goal is to use this phase sensing technique to measure the phase distortion due to the misalignment of segmented mirror so that a correction can be made by adjusting the primary mirror segments. This must be achieved notwithstanding the presence of atmospheric turbulence. The following description assumes a monochromatic system (i.e. no chromatic effects are exploited in attempting to resolve the wave-front discontinuities).

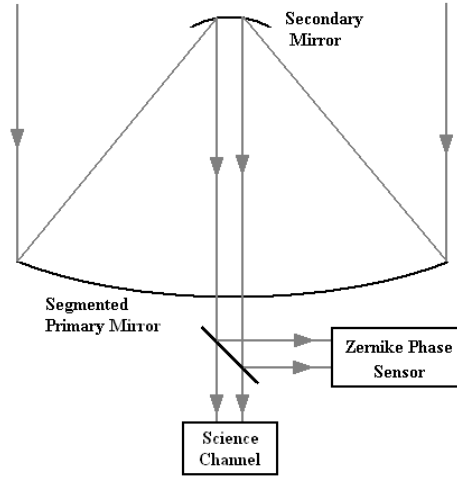


Figure 4.3 Schematic of the telescope active control system.

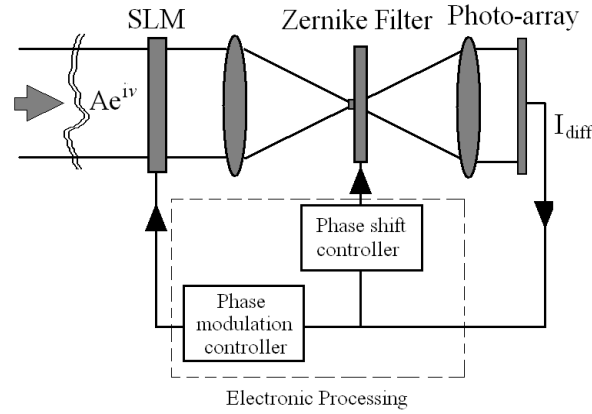


Figure 4.4 Schematic of Zernike phase sensor system; the input phase  $v$  comprises a misalignment error  $v_m$ , and atmospheric distortion error  $v_a$ .

As seen in Fig. 4.4, the differential Zernike feedback sensor comprises a spatial light modulation (SLM) as a wavefront corrector and the differential Zernike filter as a wavefront sensor; the two are coupled by a feedback loop. Here the SLM introduces a phase modulation  $u(\vec{r}, t)$  to the distorted input wave  $A_{in}(\vec{r}, t) = A_0(\vec{r}, t) \exp[i v(\vec{r}, t)]$ , where  $v(\vec{r}, t)$  is the phase distortion of the incoming wavefront, comprising both static misalignment of the segmented mirrors,  $v_m(\vec{r})$ , and time-dependent atmospheric turbulence,  $v_a(\vec{r}, t)$ , i.e.,  $v(\vec{r}, t) = v_m(\vec{r}) + v_a(\vec{r}, t)$ . The wave after the SLM is therefore  $A_{out}(\vec{r}, t) = A_0(\vec{r}) \exp\{i[u(\vec{r}, t) + v(\vec{r}, t)]\}$ , which is used as the input to the Zernike filter. The output signal  $I_{diff}(\vec{r}, t)$  from the filter is then used to control the phase modulation

$u(\vec{r}, t)$  on the SLM. The dynamics of the phase modulation in the SLM has exactly the same form as that for Gradient-Flow control (Eq. 4.16). The discrete-time version of Eq. 4.16 corresponds to the following iterative wave-front correction algorithm:

$$u^{(n+1)}(\vec{r}) = u^{(n)}(\vec{r}) + d\nabla^2 u^{(n)}(\vec{r}) - KI_{diff}^{(n)}(\vec{r}) - \mu[\bar{u}^{(n)} - u_0] \quad (4.17)$$

where  $n = 1, 2, 3, \dots$  is the iteration number,  $u^{(n)}(\vec{r})$  is the phase modulation at the  $n$ th iteration and  $I_{diff}^{(n)}(\vec{r})$  is the output signal of the differential Zernike filter photo array.

$u^{(1)}(\vec{r}) = 0$ ,  $\bar{u}^{(n)} = S^{-1} \int u^{(n)}(\vec{r}) d^2\vec{r}$  is the phase averaged over the aperture area  $S$ . The parameter  $d$  is a ‘diffusion coefficient’ that describes spatial coupling in the SLM and  $K$  is the gain parameter that can be controlled electronically in the feedback loop. We assume  $K$  to be a constant. In practice, this model is valid when the bandwidth of the incoming signal is much smaller than the central operation wavelength of the telescope. For example, for a narrowband of  $20nm$  in bandwidth and central operation wavelength at  $900nm$ , which is used in Keck telescope, the ratio of the bandwidth to the operation wavelength is less than 2% and our model works well. The last term in Eq. 4.17 is used to compensate the bias of  $u$  so that its average value goes to  $u_0$  and the coefficient  $\mu$  controls the speed of the compensation. We note that the main light loss in the Zernike phase sensor is due to SLM. We note in our system schematic figures, we show transmitting SLM only for simplicity. However, in practice, a reflective SLM is more suitable because reflective SLM has higher photon efficiency (up to 80%). SLM usually has a minimal cross-talk between neighbouring pixels justifying the assumption  $d = 0$ . We assume that both the discrete wave-front corrector and the Zernike filter photo array are matched in the sense that they have the same number of pixels and pixel geometry. In accordance with Eq. 4.17 the output signals from a differential Zernike filter photo array (after scaling by the factor  $K$  and dc component subtraction) are directly (point-to-point) mapped to the wave-front corrector array signals. This controller can be integrated with the differential Zernike filter imaging sensor, providing feedback control computation directly on the imager chip.



#### 4.4.2 Small Phase Distortion Approximation

The mechanism of the system responsible for phasing in the presence of atmospheric turbulence can be clearly understood from the small signal analysis. In the weak phase limit ( $|u| \ll 1, |v| \ll 1$ ), Eq. 4.17 with  $d = 0$ ,  $\mu = 0$ , and  $u^{(1)} = 0$ , becomes:

$$\begin{aligned} u^{(n+1)}(\vec{r}) &= u^{(n-1)}(\vec{r}) - KI_{diff}^{(n-1)}(\vec{r}) - KI_{diff}^{(n)}(\vec{r}) \\ &= u^{(n-2)}(\vec{r}) - KI_{diff}^{(n-2)}(\vec{r}) - KI_{diff}^{(n-1)}(\vec{r}) - KI_{diff}^{(n)}(\vec{r}) \\ &= \dots \\ &= -K[I_{diff}^{(1)}(\vec{r}) + I_{diff}^{(2)}(\vec{r}) + \dots + I_{diff}^{(n)}(\vec{r})], \end{aligned} \quad (4.18)$$

which shows that the time-averaged phase is an accumulation of successive iterations of the outputs of the Zernike filter. In the weak signal limit, we have  $\sin(v) \approx v$ , so Eq. 4.7 can be expressed as a linear sum of the contributions from the static and dynamical phase modulations

$$\begin{aligned} I_{diff}^{(n)}(\vec{r}) &\approx \alpha[u^{(n)}(\vec{r}) + v^{(n)}(\vec{r})] \\ &= \alpha[u_m^{(n)}(\vec{r}) + v_m^{(n)}(\vec{r})] + \alpha[u_a^{(n)}(\vec{r}) + v_a^{(n)}(\vec{r})], \end{aligned} \quad (4.19)$$

where  $u^{(n)}(\vec{r}) = u_m^{(n)}(\vec{r}) + u_a^{(n)}(\vec{r})$ ,  $v^{(n)}(\vec{r}) = v_m^{(n)}(\vec{r}) + v_a^{(n)}(\vec{r})$ ,  $\alpha = 2A_0 / \overline{A_0}$ . In deriving Eq. 4.19 we have set  $\Delta = 0$  in Eq. 4.7, which is valid under the assumptions that the amplitude of the input field is uniform and the phase fluctuations spatially averaged over the entire pupil is zero. This condition is also used in other work [30]. We note that in general the averaged phase  $\Delta$  can be an arbitrary constant and, from Eq. 4.7, setting  $\Delta = 0$  merely removes the overall piston term, which does not affect the performance of the system. Combining Eq. 4.18 with 4.19 leads to the following two equations that describe separately the feedback signals due to misalignment and atmospheric turbulence,

$$u_m^{(n+1)}(\vec{r}) = [(1 - K\alpha)^n - 1]v_m(\vec{r}) \quad (4.20)$$

$$u_a^{(n+1)}(\vec{r}) = -K\alpha[(1 - K\alpha)^{n-1}v_a^{(1)} + (1 - K\alpha)^{n-2}v_a^{(2)} + \dots + (1 - K\alpha)v_a^{(n-1)} + v_a^{(n)}] \quad (4.21)$$

where  $n = 1, 2, 3, \dots$ ,  $u_m^{(1)}(\vec{r}) = u_a^{(1)}(\vec{r}) = 0$ . The different forms of these two equations are due to different dynamics of the two phase distortion components, and this is why they can be separated. Here we have already taken into account the fact that  $v_m^n$  is fixed but

$v_a^n$  is uncorrelated for different iteration steps  $n$ . Eq. 4.20 gives the known result in the absence of atmospheric turbulence; for  $K\alpha < 1$ ,  $u_m^{(n)}$  accumulates with the increase of  $n$  and eventually converges to  $-v_m$ , which is the signal to be retrieved. However, since  $v_a^{(n)}$  is uncorrelated between different iteration frames,  $u_a^{(n)}$  does not accumulate in the same way as  $u_m^{(n)}$ . Because of the progressive weighting, the contribution to  $u_a^{(n)}$  in Eq. 4.21 comes mainly from the last few terms in the polynomial, the number of terms that should be included depend on the value of  $K\alpha (< 1)$ . We have calculated Eq. 4.21 for different values of  $K\alpha$  for sufficiently large  $n$ , in which  $v_a^{(n)}$  is generated by the Kolmogorov model [42-44] with a Fried parameter of 0.2m. The atmospheric phase errors generated vary randomly at every iteration, i.e., no frame to frame correlation in the turbulence-induced errors. The normalised standard deviation of  $u_a^{(n)}$ ,  $\sigma^* = \sigma / \overline{\sigma_0}$ , where  $\overline{\sigma_0}$  is the standard deviation of the atmospheric distortion  $v_a^{(n)}$  averaged over time, is shown to decrease monotonically with the decrease of  $K\alpha$ , as given in Fig. 4.5. The total phase signal in the feedback loop is  $u^{(n)} = -v_m + u_a^{(n)}$  for sufficiently large  $n$ . Therefore,  $u_a^{(n)}$  can be regarded as noise background on the retrieved misalignment phase. The noise represents imperfectly-averaged atmospheric perturbations, the amplitude of its residual error decreases with  $K\alpha$ . In a practical system,  $\alpha$  is usually fixed and  $K$  can be varied as the electronic gain parameter. However,  $K$  also controls the convergence rate of the system as we will discuss later; larger  $K$  results in faster convergence. So a balance between residual atmospheric errors and convergent rate should be considered in practical applications.

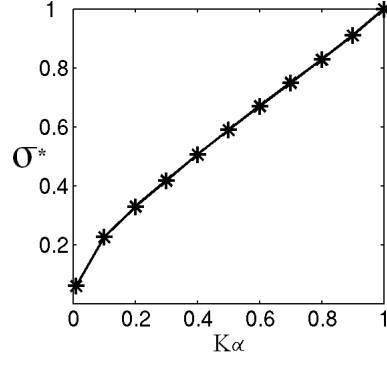


Figure 4.5 Normalized standard deviation of  $u_a^{(n)}$  as a function of  $K\alpha$ .

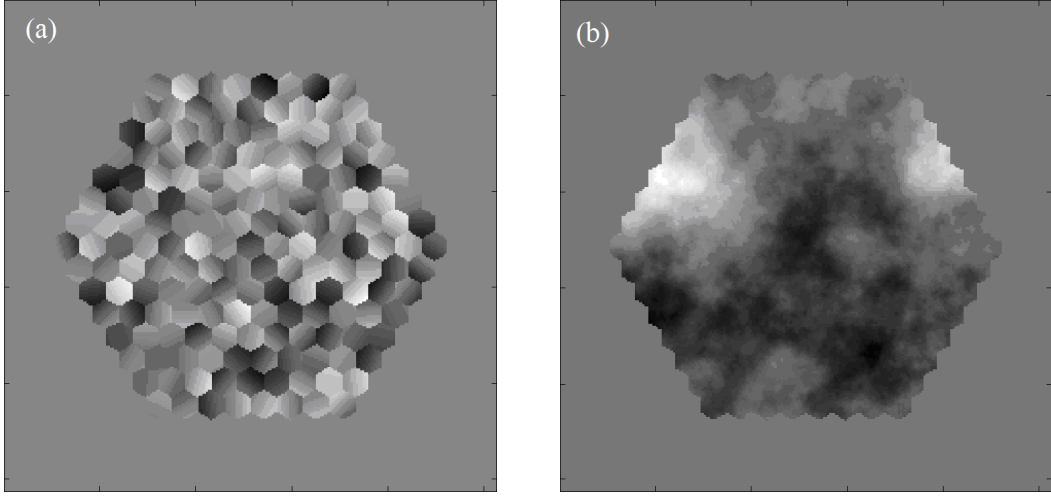


Figure 4.6 A snap shot of  $v_m$  and  $v_a$  is shown in (a) and (b), where the grey area represents the size of the SLM.

#### 4.4.3 Simulation Results

The analysis given thus far has assumed a weak phase approximation. In practice, the atmospheric turbulence will not be confined to this limit and we have explored through simulation what happens when the weak-phase approximation is violated. Our numerical findings support the above weak signal analysis.

The mirror misalignment signal is computed using a 169-hexagon formation, each of which has a random piston and tip-tilt, a snap shot of which is shown in Fig. 4.6 (a). If each mirror segment is 0.5m in diameter, the primary mirror in this simulation is 7.5m diameter. Fig. 4.6 (a) also shows the shape and size of the SLM in relation to the mirror segments. We have used  $256 \times 256$  pixels in square shape in the simulation of the SLM,

the total number of pixels overlapping with the telescope is 27288, which is much larger than the number of the telescope mirror segments. The segmented mirror should map onto the SLM for optimum use, but no particular alignment is required between the mirror segment edges and the SLM pixels. The discontinuous pixel edges in the SLM can cause scattering loss of light because of limited size of lenses in the Zernike filter. However, such lost can be kept minimal and has not been taken into account. In modelling the system we have considered the cut-off frequency due to limited spatial resolution of the SLM. We note that the alignment of the Zernike phase plate with the incoming image is a practical issue, as a displacement will result in a deviation of the phase shift spot from the centre of the Zernike focus plane and therefore reduce the processing speed of the system. Here we assume a perfect alignment in the system and the issue about the displacement will be discussed later. Atmospheric wavefront is generated by filtering the random phase with Kolmogorov spectrum [42-44] and varies randomly at every iteration, a typical spatial distribution of which is shown in Fig. 4.6 (b). A more detailed description for the atmospheric turbulence simulation is provided in the end of the thesis<sup>2</sup>. Due to low frequency inadequate sampling, low frequency aberrations, such as tilt, are underestimated and methods have been developed to compensate the low frequency component [45]. However, our system is insensitive to the exact statistics of the phase error, although it is sensitive to long-time correlation in atmospheric phase errors. The compensation is therefore not necessary in our simulation. We have not included wavefront tilt in atmospheric turbulence, as tip-tilt can be pre-compensated by tip-tilt mirror [46]. We note that to validly use the model described by Eq. 4.8, the iteration time interval between the frames should exceed the atmospheric turbulence correlation time, which is typically 30ms. Furthermore, since we assume that the atmospheric turbulence is frozen during an integration,  $I_{diff}(\vec{r}, t)$  is proportional to the number of photons registered in the photon detector over a period shorter than the atmospheric correlation time; a short-exposure image implies increased photon and read-out noise. We have investigated the effects of the instrument noise on the performance of the system. We find that the system is robust up to a certain noise

---

<sup>2</sup> See Appendix B.

level, which depends on the relative strengths of the misalignment distortion and atmospheric turbulence in the input signal and decreases with the increase of the atmospheric turbulence component. For example, for an input signal with  $4\pi$  misalignment distortion and  $\pi$  atmospheric turbulence (both peak to valley amplitudes), the retrieved signal has no significant deterioration in the presence of the (Gaussian) instrument noise with the strength of 40% (rms) of the input signal, and when the amplitudes of the misalignment and atmospheric signals are both  $4\pi$ , the tolerable noise level is reduced to 15%. We note that in practice one would use data frames recorded at a rate leading to significant frame to frame correlation in the atmospheric error and, in this case, the correlated frames will lead to a frame by frame improvement in photon and detector noise whilst the convergence of atmospheric fluctuations would be dictated by the number of independent atmospheric realizations within the data set. Therefore, the number of iterations used in the simulations with respect to the convergence of the algorithm indicates the minimum period over which the data is recorded. Since this paper focuses primarily on a new phasing method, some practical issues such as the time correlation of atmospheric turbulence and instrument noise will be investigated in more detail in the future work.

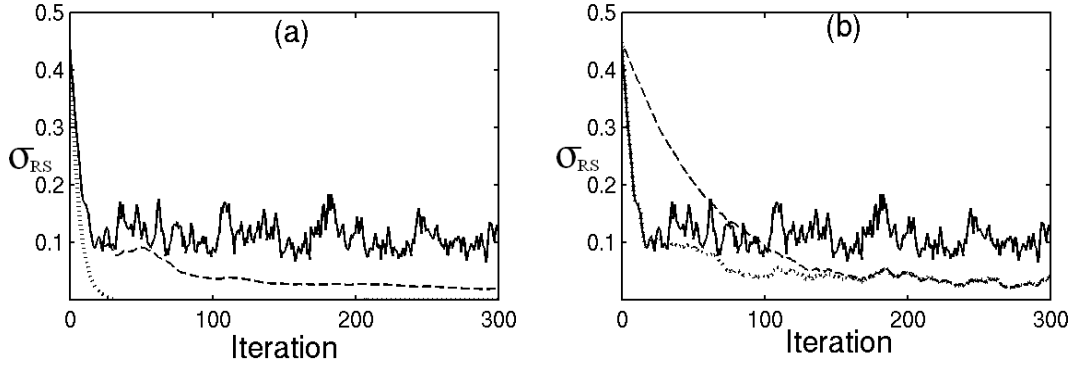


Figure 4.7 (a) The standard deviation ( $\sigma_{RS}$ ) of the retrieved signal,  $u^{(n)}$ , against the actual misalignment,  $v_m$ , as a function of the iteration number. The solid curve corresponds to  $\pm 0.4\pi$  amplitude variation for both the misalignment,  $v_m$ , and atmospheric turbulence  $v_a$ , whereas the dash curve is the standard deviation of the retrieved signals averaged over time. The dot curve corresponds to  $\sigma_{RS}$  in the absence

of atmospheric turbulence. The feedback gain is  $K = 0.75$ . (b)  $\sigma_{RS}$  for different feedback gains;  $K = 0.75$  (solid),  $K = 0.1$  (dash), and varying  $K$  (dot) : 0.75 for  $n < 20$  and 0.1 for  $n > 20$ .

Fig. 4.7 shows the result for which a peak-to-valley amplitude of  $0.8\pi$  is used for both  $v_m(\vec{r}, t)$  and  $v_a(\vec{r}, t)$ . The difference between the retrieved and original phases is measured by the standard deviation,  $\sigma_{RS} = \sqrt{\sum [(u + v_m) - (\bar{u} + \bar{v}_m)]^2}$ , where  $u$  is retrieved signal, and  $\bar{u}, \bar{v}_m$  are spatially averaged value of  $u$  and  $v_m$ . When the signal is perfectly recovered,  $u = -v_m$ , up to a constant phase shift. The term  $(\bar{u} + \bar{v}_m)$  is used to remove this shift so  $\sigma_{RS}$  gives the measure of how close the retrieved signal resembles the misalignment. As shown in Fig. 4.7 (a),  $\sigma_{RS} < 0.01$  rad is achieved for  $\sim 20$  iterations in the absence of atmospheric turbulence. When the turbulence is included,  $\sigma_{RS}$  fluctuates around  $\sigma_{RS} = 0.1$  rad. This shows a good recovery compared with  $\sigma = 0.45$  rad for the input signal, which corresponds to  $\sigma_{RS}$  at  $n = 0$  in Fig. 4.7. The asymptotic value of  $\sigma_{RS}$  is due to the residual random fluctuations of the atmospheric turbulence over a number of frames determined by the choice of  $K\alpha$ . To reduce the level of the random fluctuations we average  $u^{(n)}$  over 1000 iterations after the convergence has been achieved, i.e.,  $\langle u \rangle = \frac{1}{n-m} \sum_{i=m}^n u^{(i)}$ . We obtain  $\sigma_{RS} = 0.03$  after  $m=1000$ . We note that the presence of turbulence at this level does not appear to change the convergent rate of the system.

Fig. 4.7 (b) shows the effects of the choice of feedback parameter  $K$  on  $\sigma_{RS}$ . In general, with increase of  $K$ , the convergence rate increases, so does the noise level caused by imperfect averaging of the atmospheric turbulence. To achieve faster convergence rate and higher quality recovery at the same time, we can vary  $K$  during the iteration processing; an example is given in Fig. 4.7 (b) in which  $K$  is reduced from 0.75 to 0.1 once the convergence is achieved.

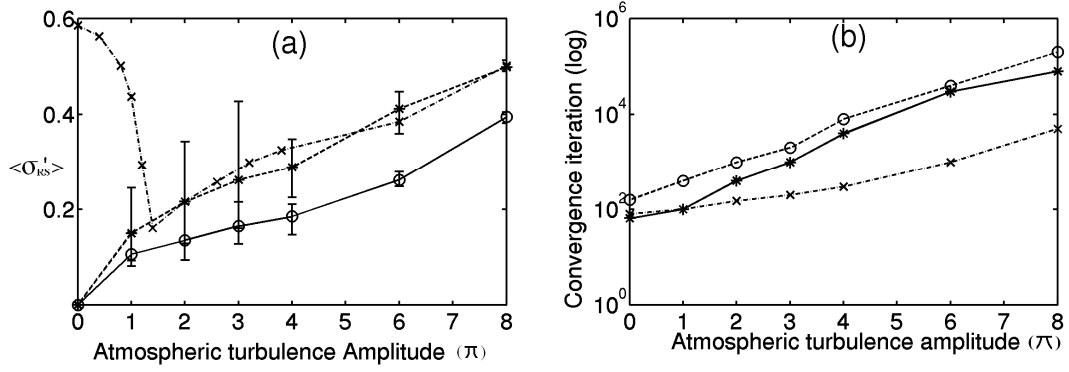


Figure 4.8 (a) standard deviation  $\langle \sigma_{RS}' \rangle$  versus different atmospheric turbulence amplitudes (peak-to-valley),  $K = 0.3$  (circle) and  $0.75$  (asterisk). (b) The number of iteration needed for convergence versus atmospheric turbulence amplitudes for  $K = 0.3$  (circle) and  $0.75$  (asterisk); by convergence it means that the retrieval misalignment signal changes very little in subsequent iterations. The curves with crosses in (a) and (b) correspond to the results of a Zernike filter with  $\pm \pi/2$  phase shift on  $5 \times 5$  pixels around the centre and for  $K = 0.75$ .

Assuming that mirror diameter is less than the outer scale, peak to valley atmospheric phase distortion increases with the diameter of telescope aperture. We have modelled here an atmospheric phase distortion of about  $8\pi$  peak-to-valley value in the absence of wavefront tilt, which corresponds to 4m primary mirror aperture diameter and Fried parameter of 20cm ( $0.5''$  seeing at visible wavelengths under the best seeing conditions). This level of error is sufficient to ensure that there is no fixed “core” speckle in the central region of the image plane and that a fully-developed speckle pattern is modelled for the snapshot point spread function. We have examined the performance of the system for  $4\pi$  (peak-to-valley) mirror-segment phase misalignment (which corresponds to the standard deviation  $\sigma = 2.24$ ) in the presence of increasing atmospheric amplitudes from 0 to  $8\pi$ . The results are evaluated by the standard deviation  $\sigma_{RS}' = \sqrt{\sum [(u' + v_m) - (\bar{u}' + \bar{v}_m)]^2}$ , where  $u'$  is the unwrapped phase in which  $2n\pi$  ambiguity is removed (see below). As  $\sigma_{RS}'$  fluctuates for each realisation, which is shown by the error bars in Fig. 4.8 (a), we average  $\sigma_{RS}'$  over 1000

iterations after the convergence of the system to obtain  $\langle \sigma_{RS}' \rangle$ . Fig. 4.8 (a) shows a monotonic (nearly linear) increase of  $\langle \sigma_{RS}' \rangle$  with the atmospheric turbulence amplitudes for both feedback coefficients  $K = 0.75$  and  $0.3$ . Consistent with the weak signal analysis, we find that  $\langle \sigma_{RS}' \rangle$  is smaller for smaller values of  $K$ . The standard deviation  $\langle \sigma_{RS}' \rangle$  in the presence of atmospheric turbulence of  $\pi$ ,  $4\pi$  and  $8\pi$  amplitudes corresponds to a measurement accuracy of 11nm, 21nm and 57nm at the operation wavelength of 900nm. Fig. 4.8 (b) shows an exponential increase of the numbers of iterations required for the system convergence with the increase of the atmospheric turbulence amplitudes. It further shows that a greater number of iterations is needed for smaller values of the feedback coefficient  $K$  (comparing circles with asterisks), which agrees with the weak signal analysis. Fig. 4.9 (a) is an example of a  $4\pi$  mirror segment misalignment phase error in the absence of turbulence-induced errors whereas Fig. 4.9 (b) is the retrieved mirror-segment misalignment in the presence of  $8\pi$  atmospheric turbulence. A clear correspondence between the two images confirms the capability of the system.

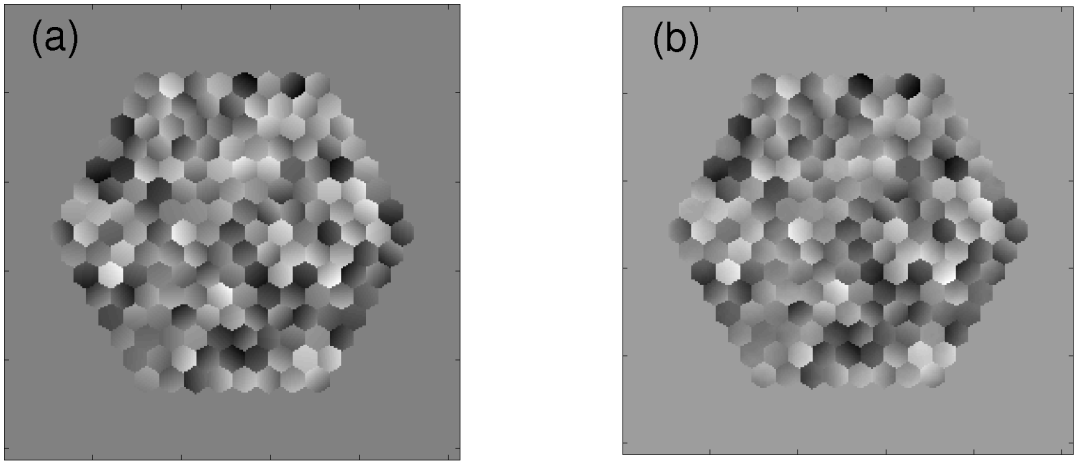


Figure 4.9 (a) Input phase distribution of a segmented mirror with peak-to-valley amplitude of  $4\pi$  misalignment. This misalignment is mixed with  $8\pi$  atmospheric turbulence as the input signal to the Zernike sensor. (b) Retrieved misalignment phase image (unwrapped), which shows very good recovery.

In our system, when the input phase fluctuations exceed  $2\pi$ , the retrieved phase images,  $u$ , are subject to  $\pm 2n\pi$  ( $n$  integer) jumps, i.e., phase wrapping. The phase



wrapping occurs because of the use of sine function in the Zernike filter, which is a common problem in phase sensing. The phase ambiguity can be determined by first measuring the phase errors in two different wavelengths and then applying the two wavelengths interferometry algorithm proposed by Lofdahl and Eriksson [26]. In our system the phase errors at two different wavelengths can be obtained as two separate measurements at different time when different filters are used in front of the SLM.

For realistic levels of turbulence the snapshot image from a large-aperture telescope has a negative exponential probability for the brightness of the axial speckle in the image plane. It follows that a Zernike filter that phase-shifts just the central diffraction-limited speckle has negligible effect on the recorded data, as we have already noted earlier that images in this situation contains no core speckles. This leads to a slower convergence for the algorithm if implemented with a pure Zernike filter. To overcome this problem, we have studied the Zernike filter with an enlarged phase shift region in the focal plane. The enlargement of the phase-shifting filter spot increases the effect that the filter has on the recorded data but degrades the quality of the interferometric phase-reference generated by the filter. This results in a trade-off between convergence rate and accuracy of wave-front reconstruction. Fig.4 gives the numerical results (crosses) for a differential Zernike filter in which  $\pm \pi/2$  phase shift is realised in  $5 \times 5$  pixels centred around the zero spectral component (compared with 1 pixel in the earlier study). As shown in Fig. 4.6 (b), the convergence time is significantly decreased, especially for larger atmospheric amplitudes where the decrease is of two orders of magnitude. However, the standard deviation  $\langle \sigma_{RS} \rangle$  using such a Zernike filter reveals a new feature, as shown in Fig. 4.6 (a). While  $\langle \sigma_{RS} \rangle$  is similar to that obtained by the conventional Zernike filter for relatively large atmospheric amplitudes, it behaves very differently when the amplitudes are below the value of around  $\pi$ . This is because when the atmospheric amplitudes are low the incoming images comprise mainly the static mirror-segment misalignment. This leads to the build-up of a considerable amount of low spatial frequency components in the retrieved phase images by the Zernike filter with  $5 \times 5$  pixels, which results in the increase of  $\langle \sigma_{RS} \rangle$  in the small atmospheric amplitude region. When the input phase image

comprises significant amplitude of atmospheric turbulence, the low spatial frequencies becomes less prominent and the results for the two cases with different pixel numbers are essentially the same. This study shows that for different strengths of atmospheric distortion we can use a Zernike filter that provides  $\pm\pi/2$  phase shift to different number of pixels in the focal plane to achieve both better quality phase image recovery and faster convergence rate. This enlarged phase shift area approach can also be used to deal with imperfect alignment between the incoming image and the Zernike phase plate and residual wavefront tilt, which was mentioned earlier. We have confirmed through numerical simulation that a better quality image recovery with faster convergence rate can be achieved when the enlarged phase shift area is applied to the case of a slight alignment displacement. Another approach to the problem is to use a time-dependent Zernike phase shift spot that dynamically follows the point of the maximum spatial frequency strength in the Zernike focal plane. A shortcoming of this method is the presence of a residual phase tilt.

#### 4.5 Conclusions and Discussion

In conclusion, we have studied an adaptive Zernike dynamical feedback system and shown by analysis and simulation that it provides an effective means to measure segment misalignment error in the presence of atmospheric turbulence. This system can be implemented in an optoelectronic device, which would provide a practical method for phasing of segmented mirror and other applications.

We note that the phasing technique introduced in this chapter is different from DZEUS and ZEUS method. In ZEUS and DZEUS, only a single exposure from the Zernike filter output is used to retrieve the segmentation error based on its known wavefront-to-intensity converting relation for segmented mirror correction at each time. The atmospheric turbulence is partially averaged out by relatively long time exposure in ZEUS and DZEUS. Measurements of a fixed-piston segment simulation plate using the ZEUS laboratory setup give a phasing error of  $14nm$  rms [15]. While in the differential Zernike feedback system, the atmospheric turbulence is averaged out by multiple exposures.

## **Chapter 5**

### **Automatic Particle Detection and Tracking in Living Cell**

Automated tracking and analysis of moving objects in image sequences has been one of the major fields in digital image analysis research. Automatic particle tracking has many applications in video surveillance, multimedia services, automated vehicle guidance and driver assistance, remote sensing and meteorology, and medical imaging. Moreover, automatic tracking is also a very important field in molecular biology [54, 55]. Biomolecular systems are dynamic and complicated, and it is one of the major challenges of biomedical research and pharmaceutical industries to unveil the spatial and temporal relationships of these complex systems. Results in this area can be expected to have significant social and economic impact in the near future, as they can improve human health and well-being. Studies into biomolecular dynamics generate huge amounts of image data. To be able to handle these data and to fully exploit them for describing biological processes on a quantitative level and building accurate mathematical models of biomolecular dynamic structures, computerized motion registration and analysis is becoming a necessity [56].

#### **5.1 Method for Study of Biological Molecular Dynamics**

Currently, light microscopy [57] is the most important imaging tool for recording of dynamic processes in living cells. Recently, Light microscopy has become mature enough to allow imaging of molecular complexes and even single molecules in living

cell. Apart from a great improvement in optics hardware and the development of increasingly sensitive electronic imaging sensors, a key factor was the discovery, cloning, and expression of the jellyfish green fluorescent protein (GFP). This enabled visible fluorescence to be encoded into a specific gene of interest, which, in turn, enables us to tag and optically detect virtually any protein of interest in living cells. Combined with time-lapse imaging, these developments have provided powerful tools to study the dynamic characteristics and functions of proteins in living cell [58].

## 5.2 Fundamental Problems of Automatic Tracking in Living Cell

Achieving robustness and high accuracy in tracking and motion analysis in images obtained by light microscopy is hampered by three factors. The first is the limited spatial resolution of the microscope. Even an optimally designed microscope, which to a good approximation can be modelled as a linear shift-invariant system with a finite point-spread function, suffers from diffraction. The Fraunhofer-diffraction limited PSF of a confocal microscope with circular aperture and operating under design conditions is given by [59]

$$PSF(r, z) = \left| \int_0^1 2J_0(\alpha r \rho) \exp(-2i\gamma z \rho^2) \rho d\rho \right|^2, \quad (5.1)$$

in which  $\alpha = \frac{2\pi NA}{\lambda}$  and  $\gamma = \frac{\pi NA^2}{2\lambda}$ , and where  $r = \sqrt{x^2 + y^2}$  represents the radial distance to the optical axis,  $z$  is the axial distance to the focal plane,  $i$  the imaginary unit number,  $J_0$  the zeroth-order Bessel function of the first kind,  $NA$  the numerical aperture of the objective lens, and  $\lambda$  the wavelength of the light emitted by the specimen. This function is band-limited in both the lateral (in-plane) and the axial (across-plane) direction, with radial cut-off frequencies of  $\omega_r = 2\alpha$  and  $\omega_z = 2\gamma$ , respectively.

The second factor is noise. Even if all sources of noise due to system imperfections are reduced to a minimum, the signal-to-noise ratio (SNR) is still limited because of the randomness introduced by the quantum nature of light. This randomness follows a Poisson distribution and is therefore not independent of the signal. Furthermore, in most experiments the signal has to be kept to a minimum, since high illumination rapidly

quenches fluorescence and may disrupt the cellular and molecular processes being studied. And for the very same reason, the number of images taken for a given time is usually minimized as well. As a result, both the SNR and the temporal resolution are usually quite low.

The third limiting factor is the large variability of biological image data. First of all, this high variability has to be attributed to the intrinsic heterogeneity of biomolecular systems. In addition, a lack of standardization in the acquisition protocols among studies may result in imagery of the same molecular process with quite different appearance and quality. The quality of images may not even be constant within one experiment, for example because of a degradation of the fluorescent probes over time (photo bleaching).

All these factors put high demands on the design of automated image processing techniques.

### **5.3 Overview of Particle Tracking Methods**

Computational image processing tools for automated tracking of molecules within living cells have been developed and reported for many years. The basic concepts underlying the vast majority of published methods are virtually the same. The commonly used approach to motion tracking consists of at least the following steps (see Fig. 1): pre-processing the image data, detecting individual particles for each image frame, linking particles detected at successive time points, and analyzing the results [54]. Pre-processing of the raw data, including noise reduction and spatial alignment of the successive images is usually required to considerably improve tracking results. Detection of relevant particles in the images is sometimes done by fitting a predefined model to the data. Once detected, a host of features can be computed for each particle, which may serve to divide particles into classes, if applicable. Feature values are also required for computing correspondence probabilities in the subsequent linking step. The resulting particle trajectories may be verified and, if necessary, corrected manually or using efficient spatiotemporal representations. Finally, a variety of dynamics parameters (velocity, direction, acceleration, etc.) may be computed from the tracks. We will discuss each of these steps in more details.

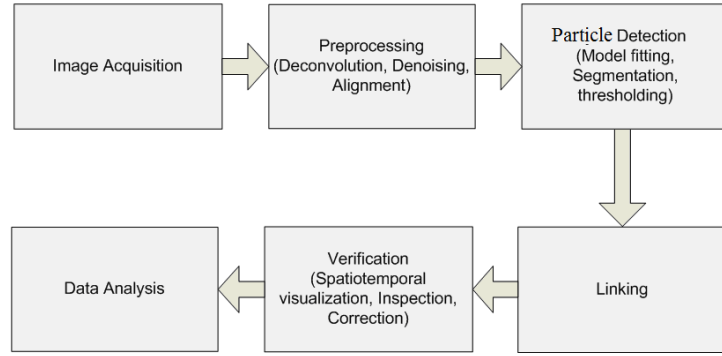


Figure 5.1 Schematic of the steps commonly involved in particle tracking studies.

### 5.3.1 Image Pre-processing

One of the most important factors influencing tracking algorithm performance is the signal-to-noise ratio (SNR). It has been demonstrated by experiments on artificial data [60] that the accuracy of commonly used tracking algorithms degrades rapidly as the SNR drops below 20 dB and becomes unacceptable below 12 dB. Such levels, however, are not uncommon in fluorescence imaging.

Moreover, it has been shown that on short time scales, localization inaccuracies caused by noise in the images may make particle diffusion processes appear anomalous even if they are normal [61]. It is therefore of crucial importance to enhance the SNR for subsequent particle tracking by applying noise reduction techniques. Since the most dominant noise source possesses Poisson rather than Gaussian characteristics, nonlinear filtering techniques are frequently used for this purpose. Examples range from simple median filtering [62] to more sophisticated anisotropic nonlinear diffusion filtering techniques [63, 64].

### 5.3.2 Particle Detection

A number of different approaches exist for estimating the positions of particles from individual images of a sequence. Most particle tracking algorithms published to date are based on either one or a combination of these methods [60]. The computationally simplest approach is to calculate the centroids, or centres of (intensity) mass, of relevant spots. This requires segmentation of the image to suppress irrelevant background structures, usually done by thresholding based on intensity or other image features.

Another, computationally more demanding approach is local image registration, where for each spot in one image the local intensity distribution serves as a template to be matched with neighbouring distributions in the next image. This requires choosing and optimizing a similarity measure, for which normalized cross correlation or the sum of (squared) intensity differences are often used. A conceptually somewhat similar, but still distinct, approach is to fit a predefined mathematical model of the spot intensity distribution. Usually this comes down to least squares fitting of a Gaussian approximation of the point-spread function.

### ***5.3.3 Particles Linking***

Once particles have been detected in all relevant frames of an image sequence, a correspondence between them needs to be established. In most practical situations this is very difficult, as the number of detected particles will generally not be constant over time. Limitations in the image acquisition process may cause not all particles to be captured at all times; particles may enter or exit the field of view, they may approach one another at distances that are no longer resolvable so that they merge into a single spot, or, conversely, a spot that seemed to represent a single particle in one frame may turn out be a cluster of particles splitting off in the next. In addition, limited detector performance at low SNR almost certainly leads to varying degrees of under- or over-segmentation (depending on the parameter settings). In combination, these factors seriously complicate the development of linking strategies.

Methods for linking corresponding particles in successive frames can roughly be classified into “local” and “global.” The former types of methods [60] operate in a per-particle fashion: each particle in one frame is linked to a particle in the next frame that minimizes a predefined distance measure. Often this involves specifying a maximum allowable distance, indicating track initiation or termination. This is the most frequently used approach to linking and may yield satisfactory results in scenes with relatively low particle densities and well-separated spots. In more complex situations with much higher densities and overlapping spots, the linking problem cannot be solved unambiguously without involving neighbouring or even all detected particles and finding the optimal correspondences for them simultaneously. Global correspondence

search strategies are well-known in image processing [65]. However, many of these perform poorly when applied to biological data because of too-simplistic assumptions of particle motion modes, which cannot cope with the intrinsic heterogeneity of motion within one particle trajectory as well as among particles. In addition, these strategies are computationally more demanding.

#### **5.3.4 Tracking Results Analysis**

Before applying a thorough quantitative analysis of the results of fully automated tracking algorithms, it is good practice to first examine and verify these qualitatively. Especially at low SNR, detection and linking errors may easily occur, even with current state-of-the-art algorithms, and require manual correction afterwards. Simply browsing through the data in a frame-by-frame fashion, as done in early studies, is cumbersome and does not provide sufficient insight into the interrelations between detected features. In the past few years, more effective ways to represent and visualize spatiotemporal data have appeared in the literature, based on (combined) volume and surface rendering techniques [56]. Apart from assisting in the verification of tracking results, such visualizations also give first impressions of possible trends in the data, which may motivate specific quantitative analyses.

Once tracking results are verified and corrected if this is possible, several characteristic motion parameters can be derived from them. Displacements, velocities, and accelerations are easily computed per particle and even per time point or interval. Generally, these values are studied collectively over larger numbers of particles, and provide the intrinsic heterogeneity of particle behaviour into histograms that reveal the most dominant modes of motion.

### **5.4 Summary**

In the last chapter of this thesis, we will show the potential application of Zernike filter in image processing, and the quantitative dynamics of a biomolecular system by a multiple particle tracking algorithm. In chapter 6, we will apply our Zernike filter feedback system, which is similar with which has been used for telescope phasing, in image pre-processing for single particle tracking. By numerical simulation, we show



this Zernike filtering can be used to remove the static background, which affects the single particle tracking process. A comparison will be given to show that the Zernike filtering will significantly improve the effectiveness of the single particle tracking algorithm. A local linking method is used in this tracking algorithm, and as a result, it can only be used in a relatively low particle density situation. Finally, we further study a global-linking multiple particle tracking algorithm and apply this algorithm to study the statistical dynamics of a bimolecular system.

## **Chapter 6**

### **Zernike Filtering for the Application of Particle Detection and Tracking**

The low signal-to-noise ratio has been a fundamental problem in dealing with particle detection and tracking in living cells. Even if all sources of noise due to system imperfections are reduced to a minimum, the signal-to-noise ratio (SNR) is still limited because of the randomness introduced by the quantum nature of light. This randomness follows a Poisson distribution and is therefore not independent of the signal. Furthermore, in most experiments the signal has to be kept to a minimum, since high illumination rapidly quenches fluorescence and may disrupt the cellular and molecular processes being studied. And for the very same reason, the number of images taken for a given time is usually minimized as well. As a result, both the SNR and the temporal resolution are usually quite low.

In this chapter, we will present our original work on further applying the Zernike filter feedback system to increase the signal-to-noise ratio of the image sequence. As we will show, this process will make the existing particle tracking algorithm more effective and accurate. The Zernike filter feedback system is capable of removing the noisy static (or slower moving by comparison with the targets) background, and therefore increasing the signal-to-noise ratio. To demonstrate the effects of the Zernike filter feedback system, we compare the tracking results with and without the presence of Zernike filtering by using numerically generated particles moving in a noisy

background. The single particle tracking algorithm we choose is a cross-correlation based algorithm, followed by centroid interpolation to give sub-pixel position. Then, we conduct case studies for the Zernike filter system in the application of living cell images pre-processing. Finally, we further present a multiple-particle tracking method on one of the real living cell image sequence, when the single particle tracking is not appropriate.

### 6.1 Zernike Filter Pre-Processing

We have used the differential Zernike filter in chapter 3 and 4 for phasing errors retrieval and correction. The differential Zernike filter had also been used in an adaptive optics system for wave-front control [29, 30]. It has been shown that the adaptive system with the differential Zernike wave-front sensor can be efficient for compensating static phase distortions. We further show in chapter 4 that this system can differentiate the dynamic component from static component in the input signal, and therefore, remove the static component.

Noisy background can hamper the effectiveness of the particle tracking in the analysis of biomolecular video images. Sometimes, there could be some static bright spots in the images which significantly affect the particle tracking. To remove the static noisy background, we can apply the Zernike filter feedback system (Zernike filtering) as a pre-processing to the image sequence to prepare the images for particle tracking.

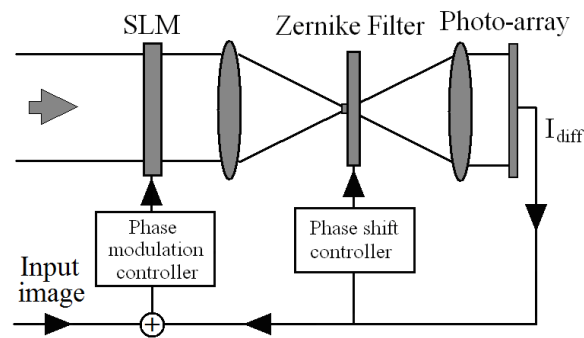


Figure 6.1 Schematic of pre-processing Zernike feedback system. The input image signal is added with the feedback signal, and then applied to the SLM.

A schematic laboratory system of the Zernike pre-processing is given in Fig. 6.1. This system, to the most extent, resembles the Zernike phasing sensor system as

described in chapter 4. The pre-processing system comprises of a spatial light modulator (SLM) as a wavefront corrector and the differential Zernike filter as a wavefront sensor; the two are coupled by a feedback loop. The only difference in this system is that the input (phase) signal is applied to the SLM rather than modulated in the input light (there is no wavefront aberration in the input wave). Here the SLM introduces a phase modulation  $u(\vec{r},t)+v(\vec{r},t)$  to the undistorted input wave  $A_0$ , where  $v(\vec{r},t)$  is the input images,  $u(\vec{r},t)$  is the feedback signal, in which,  $\vec{r}$  is the spatial radial vector in the plane transverse to the system optical axis and  $t$  is time. The wave after the SLM is therefore  $A_0 \exp\{i[\vec{u}(r,t)+\vec{v}(r,t)]\}$ , which is used as the input to the differential Zernike filter. The output signal  $I_{diff}(r,t)$  from the filter is then used to control the feedback signal  $u(r,t)$ . The dynamics of the pre-processing system can be written as a discrete-time iteration equation:

$$u^{(n+1)}(\vec{r}) = u^{(n)}(\vec{r}) - KI_{diff}^{(n)}(\vec{r}) - \mu[\bar{u}^{(n)} - u_0] \quad (6.1)$$

where  $n=1,2,3,\dots$  is the iteration number,  $u^{(1)}(\vec{r})=0$ , and  $\bar{u}^{(n)} = S^{-1} \int u^{(n)}(\vec{r}) d^2\vec{r}$  is the phase averaged over the aperture area  $S$ .  $K$  is the gain parameter that can be controlled electronically in the feedback loop. By comparison with Eq. 4.17, the diffusion term is omitted in Eq. 6.1. The output (processed) images are read directly from the photo array (known as  $I_{diff}$ ). In our experiment, this system is realized numerically. However, the hardware system can be built with the imaging camera for real-time processing to save computation for high-resolution images.

The principle of the Zernike filtering pre-processing best demonstrated through an ideal experiment is shown in Fig. 6.2. The original image sequence is constructed by adding three bright particles in a static Gaussian noisy background ( $512 \times 512$ ). The three objects have the same brightness comparable with the background, and the same FWHM of 5 pixels. Two of the objects are moving, and have constant velocity of  $V_1 = 1$  pixel/frame and  $V_2 = 0.2$  pixel/frame respectively, and the third object is static ( $V_3 = 0$ ). A snap shot of the original images is shown in Fig. 6.2 (a), and examples of

the processed images are shown in Fig. 6.2 (b), (c), and (d) respectively for frame 10, 50, and 100. We can see in the processed images, the noisy background and static object have disappeared, but the two moving objects are still there. We also note that the particle with higher speed  $V_1$  appears to be brighter than the one with relatively slower speed  $V_2$  in the processed images. This is a result of the dynamics of the Zernike filtering. The static noisy background in the initial images can be suppressed because they are built up in the feedback control signal but with opposite distribution during the iterating process. On the other hand, if an object moves so quick that this object appears in a different position at each detector reading, it can not be built up in the feedback signal, and as a result can not be suppressed. Basically, the moving objects will be completely retained while their displacement in each frame is significantly greater than their physical size. However, the object can be partly suppressed if it moves so slow that it appears to be overlapping within several iterates in the detector's reading.

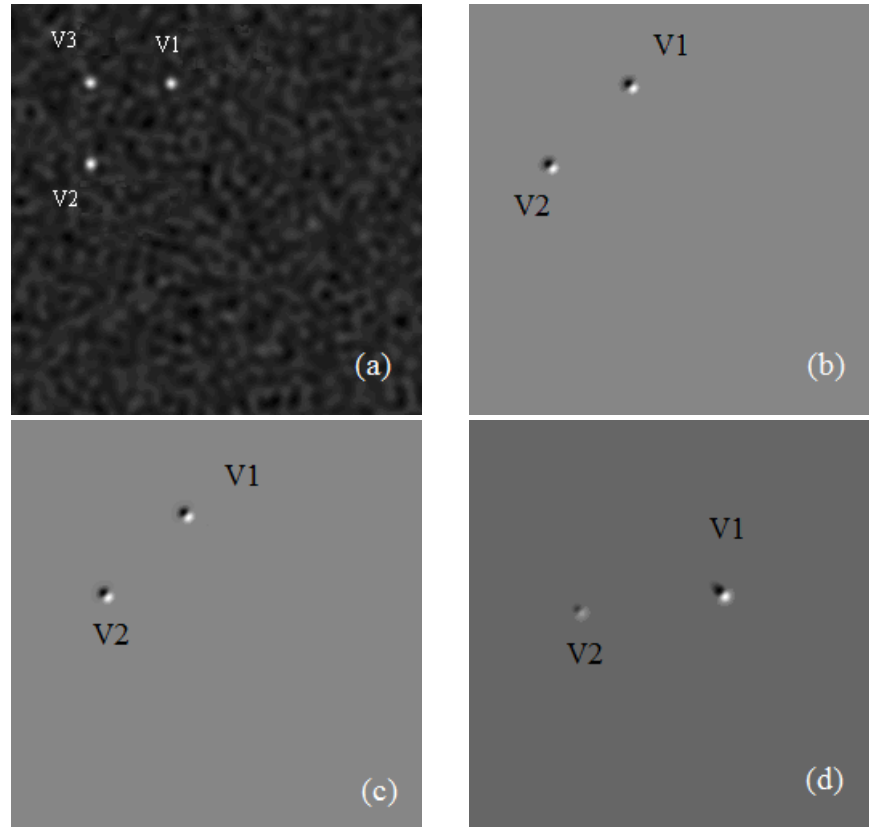


Figure 6.2 Processing results with Zernike filtering for three bright particles in a noisy background. The simulation signal comprises of three bright particles, which are corresponding to different velocity:  $V_1=1$  pixel/frame,  $V_2=0.2$  pixel/frame, and  $V_3=0$ .

(a) A snap shot at frame 1 before processing, (b), (c), and (d) are snap shots at frame 10, 50, and 100 after processing.

## 6.2 Image Pre-processing and Single Particle Tracking Procedure

As we have shown above, the Zernike filtering system is capable of suppressing the static background in images. This enhancement in image quality results in a higher signal-to-noise ratio, and therefore makes it a good candidate in pre-processing for particle tracking. In order to show the effectiveness of the Zernike filtering in this application, we use a single particle tracking method on both artificially generated and real living cell image series, and compare the tracking results with and without the presence of the Zernike filtering.

There are many methods used for single particle tracking [66], including finding the centre of mass (centroid) of the object, fitting a Gaussian curve to the object, cross-correlation (COR) and sum-absolute difference (SAD). Methods such as COR and SAD do not give sub-pixel position on their own. Instead, they create a ‘correlation image’ showing regions of high similarity between a template image and the current image. A method of interpolating the nearest sub-pixel position of this ‘correlation object’ is then necessary. The types of interpolation that have been used include parabolic, cosinusoidal, and Gaussian fitting and centroid calculation. For the method of single particle tracking, tracking begins with the selection of a region of interest. Then, in each image, the chosen method of locating the particle’s centre is applied and this centre position is used to reposition the area of interest for the next image processing. Depending on the exact implementation of particle tracking used, investigators typically propose that an accuracy of 10 nm or better can be achieved [66]. When magnification and the physical size of the pixels in the CCD chip of the camera are considered, the image pixel size is typically between 30 nm/pixel and 150 nm/pixel. Thus, to achieve an accuracy of 10 nm corresponds to locating the centre of the object to somewhere between 1/3 and 1/15th of a pixel. Currently, many choices for tracking method are to apply a COR-based algorithm, followed by centroid interpolation to give sub-pixel position [66]. This tracking method is the one we use for single particle tracking along with the Zernike filtering pre-processing. The details for this algorithm are given in

## Appendix C.

As shown in Fig. 6.3, the pre-processing and tracking procedure is implemented in the sequence of the following steps: (1) the image series are prepared or generated artificially by combining the object and noisy background, (2) the images are sent frame by frame for Zernike filtering, (3) the kernel and the coordination of the centre of searching area in the next frame are initialized, (4) the kernel and the specified searching area are used to calculate the cross-correlation coefficient, (5) a threshold operation is performed to the correlation matrix, and then centre of the mass (centroid) of the correlation matrix is calculated, (5) the object position coordinate is calculated. The process from step (3) to step (5) will be repeated until the last frame is processed.

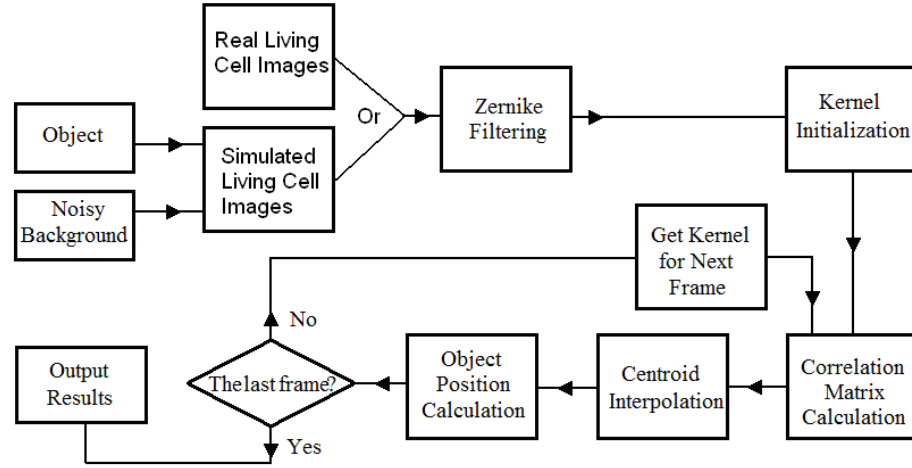


Figure 6.3 The diagram of the single particle tracking algorithm. The Zernike filtering is applied along with the particle tracking algorithm.

### 6.3 Tests on Simulated Living Cell Images

The currently existing single particle tracking methods, including the correlation combined centroid method, are capable of quantifying the position and motion of the moving objects accurately under relatively large signal-to-noise (SNR) conditions [66]. However, when the SNR decreases (typically under 10), most of these methods fell. In this section, we will show by numerical simulation that the Zernike filtering algorithm is capable of increasing the SNR by suppressing the static noisy background, and as a result, increasing the effectiveness of particle tracking methods.

We consider physical particles that are mobile in a two-dimensional plane. Their motion is observed using imaging equipment and a digital (CCD) camera which

generates a sequence of digital images at discrete time points. We call this sequence a movie and an individual image from it a frame. In each frame, the images of the particles are visible as objects.

### 6.3.1 Particle Model

To create an accurate model for a fluorescent object imaged with a charge-coupled device (CCD), the distribution of intensities of the image can be derived by convolving the object function with an appropriate point-spread function (PSF) [66]. To study the effectiveness of the Zernike filtering, we simplified this step and only use a two-dimensional Gaussian surface to represent the particle. The input image is obtained by combining the object and noisy background, which can be written as:

$$I^{(i)} = I_{obj}^{(i)} + I_{bg}^{(i)} \quad (6.2)$$

where  $I_{obj}^{(i)}$  is the object component, and  $I_{bg}^{(i)}$  is the background component, and  $i = 1, 2, 3, \dots, N$  is the frame number. The object component  $I_{obj}^{(i)}$  is given by the expression:

$$I_{obj}^{(i)}(x, y) = a \exp\{-b[(x - x_i)^2 + (y - y_i)^2]\} \quad (6.3)$$

in which  $a$  determines the brightness of the moving object, and  $b$  determine the size of the object. In our simulation  $a$  can be varied, while keeping the noisy background constant, to provide different signal-to-noise ratio, and  $b$  is set to 0.15, which makes the object of 9 pixels in diameter.  $(x_i, y_i)$  is the pre-defined position of the particle at frame  $i$ , and the predefined parabolic trajectory of the object is given by:

$$x_{i+1} = x_i + 1 \quad (6.4)$$

$$y_i = -0.011(x_i - 256)^2 + 450 \quad (6.5)$$

where  $i = 1, 2, 3, \dots, 372$ , and  $x_1 = 70$ ,  $y_1 = 69$ . The parabolic trajectory is shown in Fig. 6.4, which emerges from the left end, and exits at the right end of the curve.



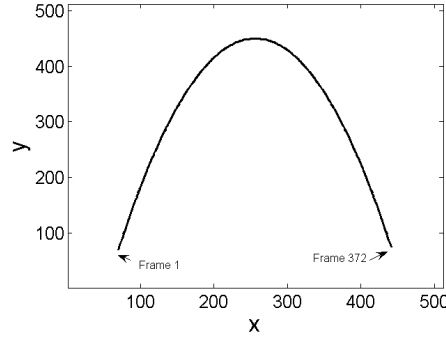


Figure 6.4 The pre-defined trajectory of the moving object.

### 6.3.2 Noise Model

The CCD cameras used for image recording generate shot noise in the image [67]. Shot noise is a Poisson process where the noise increases with  $\sqrt{N}$ ,  $N$  being the number of detected photons or photoelectrons in a pixel. In our simulation, the noisy background  $I_{bg}^{(i)}$  is realized by Poisson distributed shot noise, which does not change at each frame. If the expected number of detected photons in this interval is  $\lambda$ , then the probability that there are exactly  $k$  detected photons ( $k$  being a non-negative integer,  $k = 0, 1, 2, \dots$ ) is equal to

$$P(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (6.6)$$

where  $k!$  is the factorial of  $k$ .

### 6.3.3 Results

The artificial images, generated by combining the particle and noise, are sent for Zernike filtering before applying the tracking algorithm. We numerically choose a perfect phase shift size of one pixel [29], which physically represents shifting only the zero spatial spectral components, and the feedback gain coefficient  $K$  is chosen to be 0.2 in the pre-processing. A snap shot of the Zernike filter processed images and a snap shot of the original images are shown in Fig. 6.5. The noise background is suppressed heavily as a result of the Zernike filtering.

The numerically generated images have the size of  $512 \times 512$  pixels. In the initialization step, the centre of mass of the whole image at the first frame is used as the

centre to get the kernel and the centre of the searching area for the next frame. The template image (kernel) size is  $9 \times 9$  pixels and the search area size is  $17 \times 17$  pixels, which results in a square correlation matrix with  $17 + (9 - 1) \times 2$  pixels in each side.

The coordinates of the particle can be updated at frame  $i$  by

$$\begin{aligned} x_i &= x_{i-1} - w + x'' - w + x', \\ y_i &= y_{i-1} - w + y'' - w + y', \end{aligned} \quad (6.7)$$

where  $(x', y')$  and  $(x'', y'')$  are the centroid interpolation of the kernel (last frame) and the correlation matrix respectively, and  $w = 9$  is the width of the kernel.

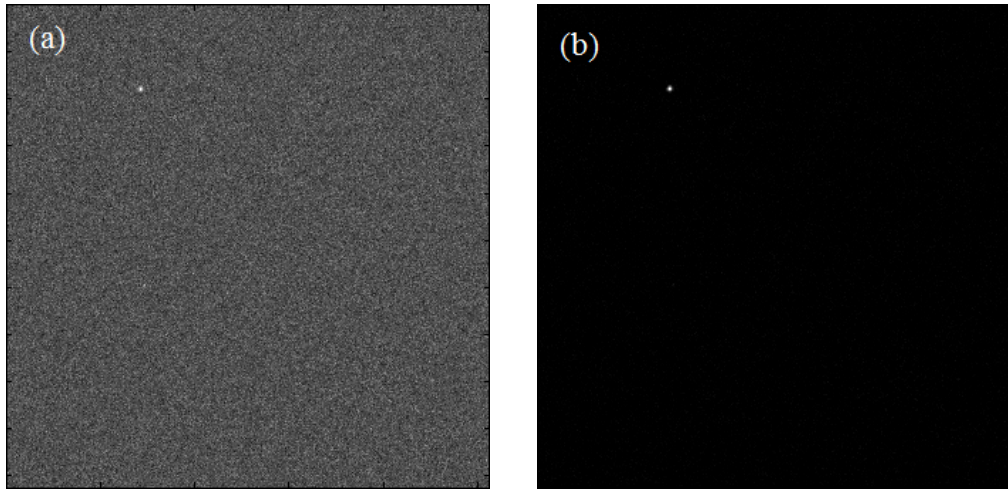


Figure 6.5 A snap shot of the original simulated images (a), and a snap shot of the Zernike filter processed images.

In order to compare the accuracy of the single particle tracking algorithm with and without Zernike filtering for different levels of relative signal intensity, the signal-to-noise ratio must be calculated. The signal-to-noise ratio (SNR) is calculated as [68]

$$SNR \approx \frac{C_{on} - n \langle x_{back} \rangle}{\sqrt{(C_{on} - n \langle x_{back} \rangle) / G + n \sigma_{back}^2 + n \sigma_{back}^2 / p}}, \quad (6.8)$$

where  $C_{on}$  is the total pixel value of a box centered on the object,  $n$  is the number of points in this box,  $x_{back}$  is the mean background from a box not centered on the object (anywhere away from the object),  $p$  is the size of this second box and  $G$  is the gain of the system.

Usually for images from a camera the gain is not known. It can be estimated, however, from the standard deviation of the background of an image. To estimate the gain, first calculate the mean and standard deviation of the pixel values in a region of the background. The noise (standard deviation,  $\sigma$ ) of a Poisson process should be  $\sqrt{N}$ , where  $N$  is the number of photoelectrons giving rise to the signal. If the gain is 1, then  $B = \sigma^2 = N$  where  $B$  is the mean background pixel value. If the gain is not one, then  $B \neq \sigma^2$  and the ratio gives the gain, thus  $G = B / \sigma^2$ .

As a result of the effect of the noise (or residue noise), there is always an error for the obtained particle coordinate at each frame in the tracked trajectories. To quantify the overall accuracy of the algorithm, the standard tracking error (SE), which measure the standard deviation of the measured object positions from the real ones of all image sequence, is introduced here

$$SE = \sqrt{\sum_{i=1}^n [(x_i - x_i^0)^2 + (y_i - y_i^0)^2] / n}, \quad (6.9)$$

where  $(x_i, y_i)$  is the measured position of the particle at frame  $i$  and  $(x_i^0, y_i^0)$  represents the real position of the particle,  $n$  is the maximum frame number.

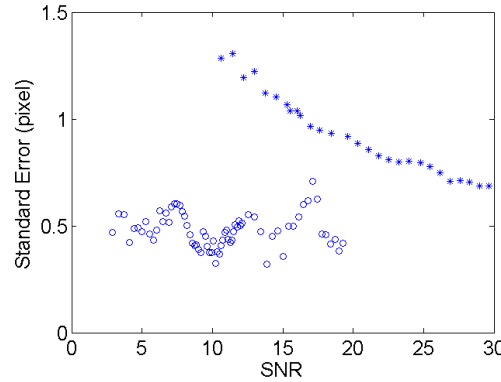


Figure 6.6 The standard tracking error as a function of signal-to-noise ratio (SNR) with (circle) and without (asterisk) the presence of the Zernike filtering. The particle tracking algorithm accuracy is quantified by the standard error (SE).

The single particle tracking algorithm works very well with the presence of pre-processing by Zernike filtering, and results in a good accuracy ( $SE < 1$  pixel). By changing the value of  $a$  in Eq. 6.6, we can change the particle's brightness and therefore the value of SNR of the image sequence, and then check the effectiveness of

the Zernike filtering. We find that the single particle tracking algorithm with Zernike filtering still provides less than 1 pixel accuracy even when SNR is as small as 3. As we know, this is better than the single particle tracking method without the Zernike filtering, which works with smallest SNR around 10.

To compare the two, we calculate the standard error of the particle tracking results with and without the presence of Zernike filtering, in the case of different SNR values. As shown in Fig. 6.6 the standard tracking error without Zernike filtering decreases when the SNR increases, while the standard tracking error with the presence of Zernike filtering does not change monotonically as what happens when Zernike filtering is not applied. This is because that most of the noisy background has been removed by the Zernike filtering and then the increase of the object brightness does not increase the accuracy of the tracking algorithm any more.

#### **6.4 Tests on Real Living Cell Images**

Three sequences of real living cell images are studied here. These case studies help us to identify the effectiveness and suitability of the Zernike filtering system in living cell imaging applications with different conditions. Fig. 6.7 (a) and 6.8 (a) are snap shots of two original fluorescence time-lapse sequences of GFP-tagged end-binding protein 1 (EB1-GFP) tracking the plus ends of extending microtubules at the posterior of *Drosophila melanogaster* mid-stage oocyte. Fig. 6.9 (a) is a snap shot of a time-lapse sequence following the localisation of fluorescently labelled mRNA in a syncitial blastoderm stage of a *Drosophila melanogaster* embryo. The corresponding snap shot after Zernike filtering are given in Fig. 6.7 (b), 6.8 (b) and 6.9 (b) respectively. The static bright area in original images is removed by Zernike filter in the processed images. We note that the significance of Poisson noise in these three living cell image sequences is descending with the increase of the image sequence number (6.7, 6.8, and 6.9), and in Fig. 6.9 there is little Poisson noise. It is not hard to find out that, the particles are mixed with residue noise in Fig. 6.7 (b) and 6.8 (b), and in Fig. 6.9 (b) the particles are basically free from noise pollution. This difference can be ascribed to the characteristics of the Zernike filter that it can not suppress the dynamic Poisson noise, but more suitable for removing static background.

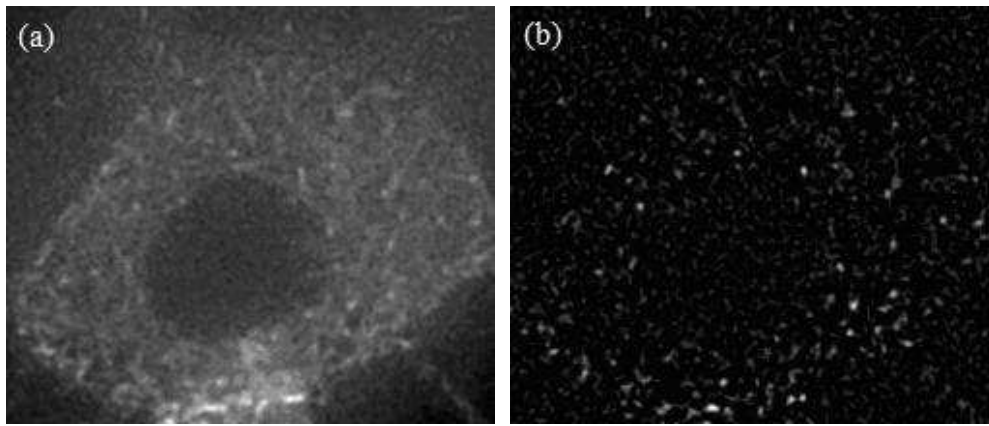


Figure 6.7 A snap shot of the original *Drosophila melanogaster* mid-stage oocyte image sequence (a) and corresponding Zernike filter processed image (b).

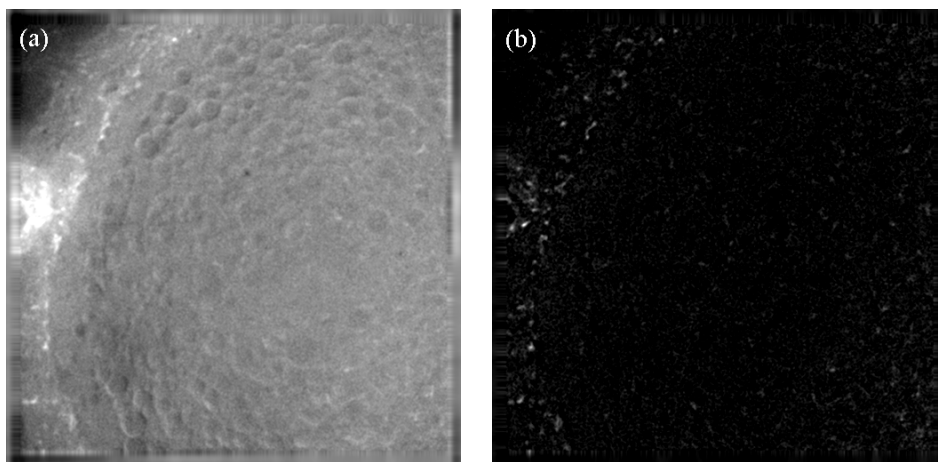


Figure 6.8 A snap shot of another original *Drosophila melanogaster* mid-stage oocyte image sequence (a) and corresponding Zernike filter processed image (b).

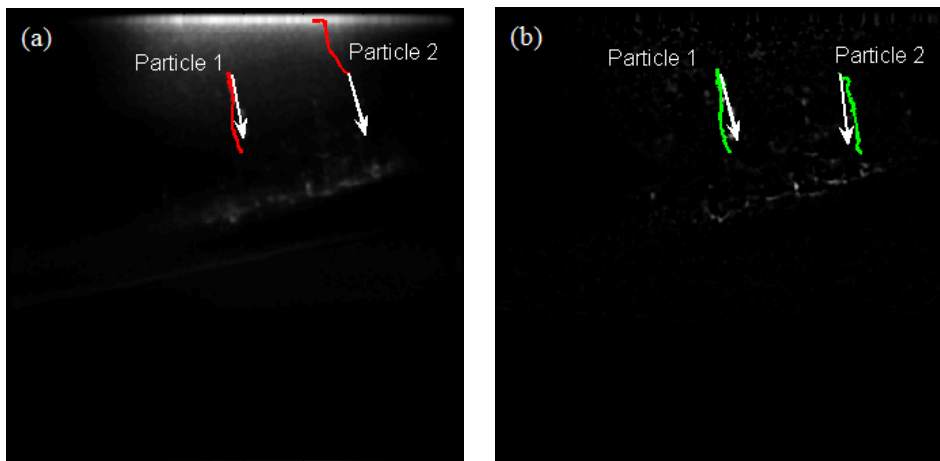


Figure 6.9 (a) A snap shot of the original *Drosophila melanogaster* embryo time-lapse sequence with fluorescently labelled mRNA, the red lines are tracks obtained without Zernike filter; (b) A snap shot of the image sequence after Zernike filter processing, the green lines are tracks obtained with Zernike filter.

As the last image sequence (Fig. 6.9) has lower particle density, we are able to conduct the single particle tracking algorithm on it. In Fig. 6.9 (a), the trajectories in red obtained by the single particle tracking algorithm without the process of Zernike filter. The particle 1 is relatively bright, and can be tracked successfully. However, the much weaker particle 2 is not successfully tracked (the arrow indicates the right moving direction). By comparison, when Zernike filter is applied (as shown in Fig. 6.9 (b)), both particle 1 and 2 are successfully tracked (the green lines). We further strengthened our conclusion that Zernike filter can make the single particle tracking algorithm working in worse conditions (low SNR).

## **6.5 Multiple-Particle Tracking**

It is easy to find that in both Fig. 6.7 and 6.8 the particle density is high, and then the linking problem cannot be solved by the cross-correlation method locally without involving neighbouring. Therefore, a global linking method is needed to find the right correspondences for all particles recognised in each frame. Global correspondence search strategies are well-known in image processing [65]. However, many of these perform poorly when applied to biological data because of over-simplistic (mostly global) assumptions of particle motion modes, which cannot cope with the intrinsic heterogeneity of motion within one particle trajectory as well as among particles. Most of these methods fail to deal with the common situation in biology when the objects enter or exit in the vision.

In this section, we employ an existing computationally efficient, two-dimensional, multiple-particle tracking algorithm [69], on the *Drosophila* oocyte image sequence (Fig. 6.7). The particles' motion in this image sequence is complex, and particles' spatial feature varies a lot from time to time. They cross each other, enter and exit in the vision. All these characteristics make it not easy to deal with by most existing tracking methods. To solve these problems, a global linking algorithm employing a graph theory technique [72-74] to determine optimal associations between two time points is used here.

### **6.5.1 The Tracking Algorithm**

The multiple-particle tracking algorithm has to perform three distinct steps. Firstly it has to conduct a pre-processing to remove the harmful noise. Secondly it will detect the particles in every frame. Finally it has to link these target detections into trajectories. Most of the details about the particles detection and linking method are omitted [69]. Here we only provide the detail about the difference between the Zernike filter and the Gaussian-like filter (for pre-processing) used in Ref. 69, and the different cost function chosen in linking procedure used in our study.

As the Poisson noise in *Drosophila* oocyte images is significant, the Zernike filter pre-processing is no longer suitable. A Gaussian-like filter has been used as a starting point of the work by Crocker and Grier for the detection of gold colloids in micrographs [76]. This filter is effective for suppressing Poisson noise. A snap shot of the Gaussian-like filter processed image sequence is given in Fig. 6.10 (b). To compare it with the Zernike filter, which was used in the section 6.3 and 6.4, the typical SNR value of the particle after processing for two images series (indicated in Fig. 6.7 and 6.8) with a different pre-processing algorithm can be calculated. The SNR for the image indicated in Fig. 6.7 (the same as Fig. 6.10) is 5.6 for Zernike filter while 8.5 for the Gaussian-like filter after processing. For the image shown in Fig. 6.8, the SNR is 6.3 for Zernike filter and 4.3 for Gaussian-like filter after processing. Better performance is attributed to the significance of either static background or dynamic noise. When the static bright background dominate, Zernike filter is better. As the dynamic noise is more significant in the image series we are working on, we choose the Gaussian-like filter for this sequence.

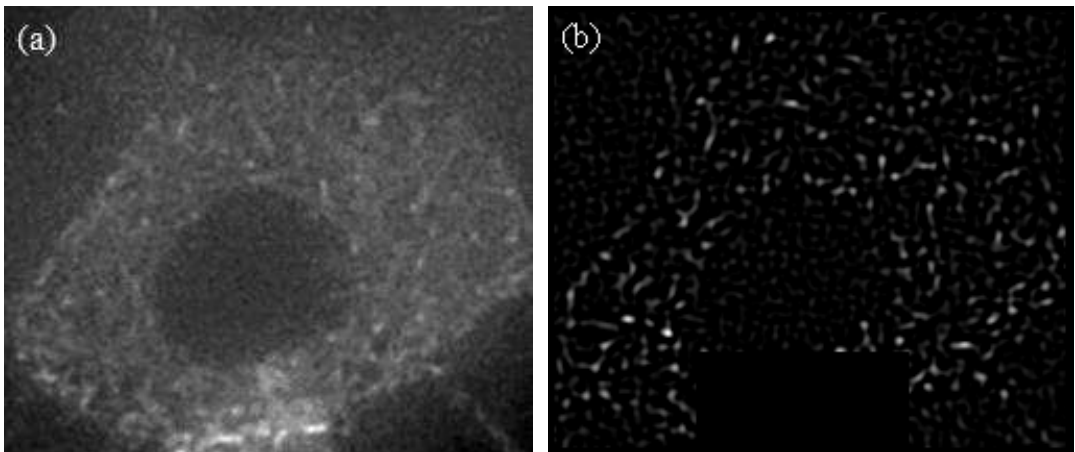


Figure 6.10 A snap shot of the original oocyte image within the interested area (a) and a snap shot of the restored image (b).

The particle association cost function is calculated to obtain the optimum global particle linking. In Sbalzarini's work [69], he used the quadratic distance between  $p_i$ ,  $i > 0$ , and  $q_j$ ,  $j > 0$ , as well as the quadratic differences in the intensity moments of order 0 and 2. As we can see in Fig. 6.11, the frame-by-frame change in second order intensity moment for a sample manually tracked particle is not stable. This is a result of significant changes in the particles' shape. Therefore, in this case, we only use the quadratic distance and quadratic difference in  $m_0$ , but not  $m_2$ .

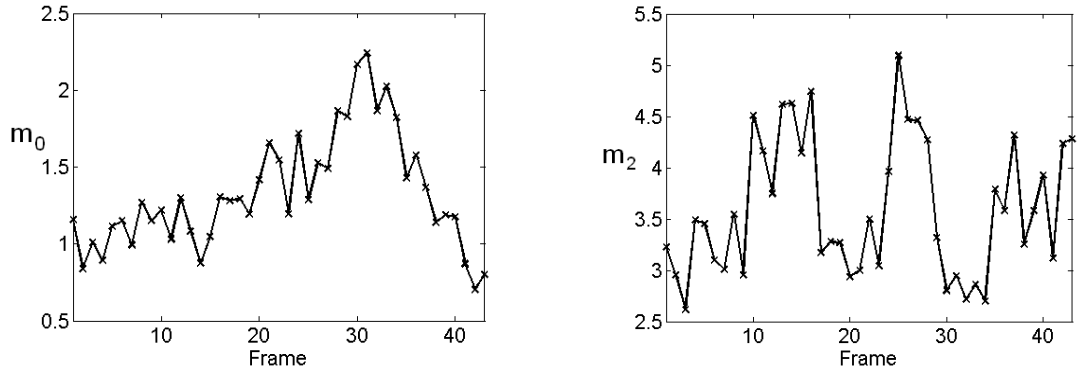


Figure 6.11 The intensity moment of order 0 ( $m_0$ ) and order 2 ( $m_2$ ) for a sample manually tracked particle evolving with time. There is a good continuity in  $m_0$ , but not in  $m_2$ .

### 6.5.2 Tracking Results and Comparison with Manual Tracking

Before presenting the automatic tracking results, we perform a manual tracking on part of the image sequence. Manual tracking is labour intensive, costly, inaccurate, and poorly reproducible, and usually only a small fraction of the data can be analyzed in this manner. However, as the statistical characteristics of the particles in the image sequence is not clear, we need to provide manual tracking to choose parameters for the automatic multiple-particle tracking. The *Drosophila* oocyte image sequence comprises 198 frames of  $512 \times 512$  pixel images. We manually tracked as many of the particles as we



could, in a  $32 \times 32$  square area, whose top left-hand corner is (352,192), and found 21 trajectories in this area.

The manually tracked coordinate data for a typical particle is shown in Fig. 6.12. This particle enters in frame 99, and disappears in frame 138. The coordinate and instant velocity of this target in axis  $x$  and  $y$  are given separately in Fig. 6.12 (a) and (b). We find this particle has an overall orientation tendency. However, the instant velocity is still somehow random. This motion can be seen as a biased Brownian motion. Another fact we find here is that the maximum instant velocity in either direction ( $x$  or  $y$ ) does not exceed 3pixels/frame. Therefore, the overall frame-by-frame displacement will not exceed  $3 \times \sqrt{2} = 4.24$  pixels between two successive frames.

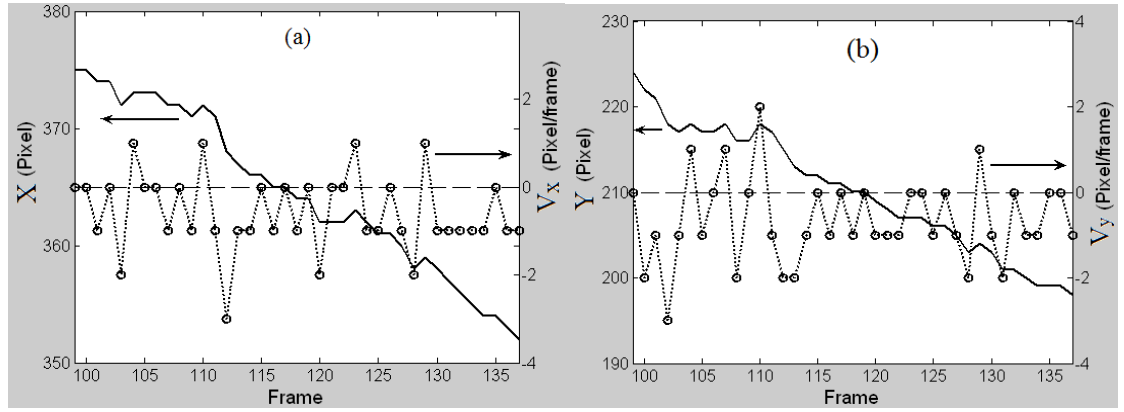


Figure 6.12 Manually tracked particle coordinate in  $x$  (a) and  $y$  (b) axis. The instant velocity in  $x$  and  $y$  axis is also shown in (a) and (b) respectively.

Another parameter needed to be properly chosen is the threshold intensity percentile value ( $r$ ) used in target location estimation. If the value of  $r$  is too small, the tracking algorithm could miss quite a few particles in certain frames. On the other hand, if  $r$  is too large, some relatively brighter noise or background spots could be treated as particles. To find the optimum value for the threshold intensity percentile  $r$ , we execute the tracking algorithm with different value of  $r$ , and find 3 percent to be the optimum value of  $r$ , which results in the best match between the manually and automatically tracked results. We run the multiple-particle tracking algorithm on the image sequence, and find 3550 particle trajectories. We note that there are many trajectories that only last for few frames. These “short” trajectories do not provide much

information for us and can be ignored. The trajectories of particles with more than 8 pixels displacement are plotted in Fig. 6.13 (a), and the trajectories of particles with more than 5 pixels and less than 8 pixels displacement are plotted in Fig. 6.13 (b). In Fig. 6.13, different colours indicate different trajectories. Other parameters used for tracking in the *Drosophila* oocyte images are: particle radius  $w=3$  pixel; noise correlation length  $\lambda_n=1$  pixel.

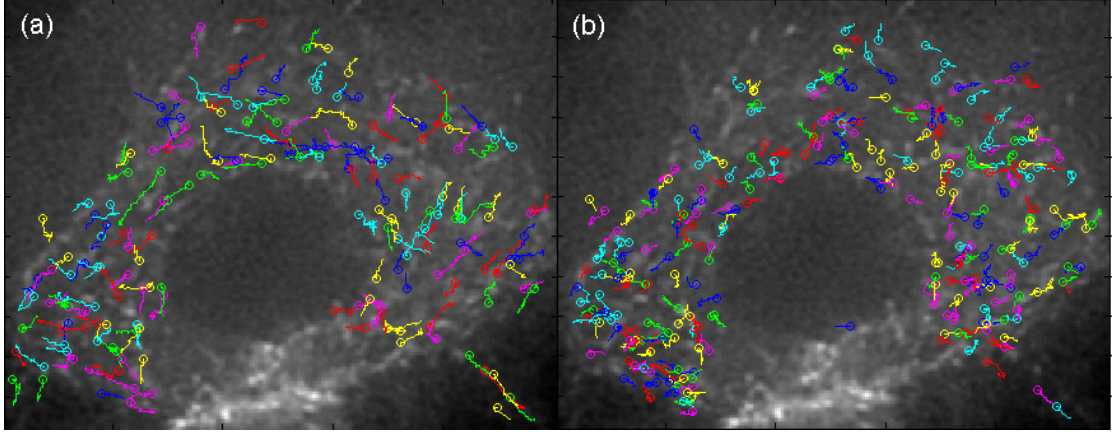


Figure 6.13 Particle trajectories resulting from the multiple-particle tracking algorithm. The different colours indicate different trajectories. (a) 187 trajectories of particle with more than 8 pixel displacement, (b) 287 trajectories of particle with more than 5 pixel and less than 8 pixel displacement.

To find out how good this tracking algorithm is, we compare the automatic tracking results with the manually tracked results. In the  $32 \times 32$  square area, the automatic tracking algorithm has generated 105 particle-linking in first 20 frames, while 11 of them are missing by comparison with manually tracking results, and no false linking exists. The linking success rate is 89.5% by comparison with manually tracked trajectories. The missed link is a result of the unstable intensity of the particles. Some particles can be very weak in certain frames, such that the tracking algorithm does not treat them as particles as their maximum intensity is lower than the threshold value.

To further verify the maximum displacement between two frames for the particles does not exceed 4 pixels, we have also run the multiple-particle tracking algorithm with different value of maximum linking length  $L=3,4,5$ , and 6 pixels. The automatically tracked results with different values of  $L$  can be compared with the manually tracked

results, and we find the optimum maximum linking length is  $L = 4$  pixels, which gives us the best balance between less false link and less unwanted trajectory break up.

The above comparison between the automatic tracking algorithm and manually tracked results shows a fair consistency between the two. However, the problem in the automatic tracking method is that it missed a few targets in several frames because they are too “weak”. These “weak” particles found in manual tracking benefit from the advantages of the human vision system. Although, at this stage, the automatic tracking algorithm can not find all the trajectories accurately, the tracking results are still valuable statistically when manual tracking is extremely difficult for such a particle density.

## 6.6 Conclusions and Discussion

In conclusion, we have shown by both simulation and case study that the Zernike filtering is capable of suppressing the static noisy background for images, which makes it a good candidate for preparing images for particle tracking (pre-processing). We show that a single particle tracking algorithm works even with relatively weak signal (SNR as small as 3) with the presence of Zernike filtering. As long as the particle density is low enough so that separate tracking of each particle is possible, the single particle tracking algorithm (a local linking method is employed) along with the Zernike filtering works effectively. However, in more complex situations with much higher particle densities and overlapping spots, the linking problem cannot be solved locally. Therefore, a global linking method is needed to find the right correspondences for all particles recognised in each frame. As an example, we applied an existing computational efficient and robust algorithm, in which a global linking method is employed, for two-dimensional multiple-particle tracking on a *Drosophila* oocyte image sequence. This algorithm works fairly well (with 89.5% linking successful rate, and no false link). We show, by comparison with manual tracking, this algorithm is limited as the tracking fails when particle intensity is too low.

We note that the Poisson noise in our single particle tracking simulation does not change with time. This is not the case in the real application. In this chapter, we use the Poisson noise only as an example of the noisy background (not for short noise). In the

real living cell images, sometimes, there are some static bright objects or illumination areas which are not of interest. In such cases, Zernike filtering can remove these irrelevant bright spots or areas and makes the particle tracking algorithm work more effectively and accurately (example shown in Fig. 6.7). However, when the background is moving, or the Poisson noise (dynamic) is significant, we have to be very careful in using Zernike filtering (example shown in Fig. 6.6), as it may also bring in extra noise. As we show in section 6.5, when Zernike filtering is not appropriate, other pre-processing method will be applied.

One problem for the Zernike filtering is that the static noisy background in the first few frames isn't suppressed as much as in the rest of the frames. This is a result of the Zernike feedback system dynamics. The static components in the images can be suppressed because they are built up in the feedback control signal during the process. However, in processing the first few frames, the feedback signal has not yet been built completely, so the static components in these frames do not disappear completely. This defect sometimes can make the tracking accuracy different in the first few frames from the rest. To solve this problem, we can artificially add the first few frames of the image series before the original image series in the opposite sequence, which generate a new image series. The Zernike filtering can then be applied to the new image series, and after the processing, the artificially added frames can be ignored in the tracking process. The added frames are used to build up the static components in the feedback signal before processing the 'real' frames.

We note the automatic multiple-particle tracking algorithm is limited in terms of accuracy by comparison with manual tracking, as it is not capable of recognizing "weak" particles. However, this is not the only limitation for this multiple-particle tracking algorithm. There are more situations in which the linking between frames is incorrect. Two examples of possible false linking are shown in Fig. 7.8. In Fig. 7.8 (a), the three moving particles are correctly linked. In Fig. 7.8 (b), there is a missing measurement. This defect does not affect the statistic characteristics of the particles significantly. In Fig. 7.8 (c), as there is an extra or false recognized particle, the link could lead to a false trajectory.

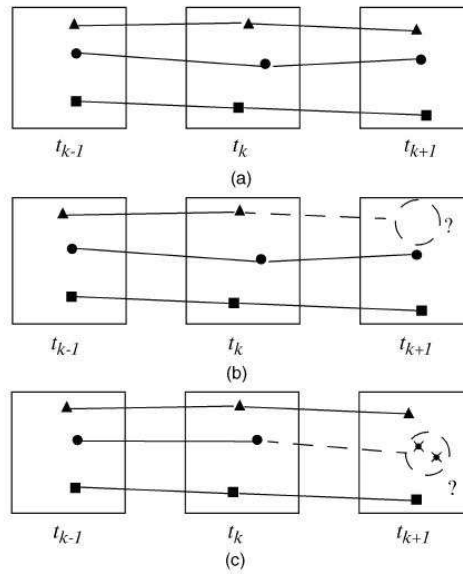


Figure 7.8 Three moving particles are measured in three time instances. The lines represent the particle correspondences in time. In (a), all points are measured at every time instance (frame). In (b), there is a missing measurement at  $t_{k+1}$ , and, in (c), there is an extra or false measurement at  $t_{k+1}$ .

One way to find and correct the false linking is to check the trajectory results. In example shown in Fig. 6.13, we know in advance that the particles move in smooth trajectories (in the long time scale), if we find there is a significant change of direction in the particle trajectory, this indicates that a false link may exist, which may then be manually corrected. Alternatively, linking to the frame after the next can be conducted in addition to the one frame linking. The results of the two can be compared, and if there is a contradiction, we can go back to the video, check the original images and correct the false link if there is one.

Advances in biological imaging technology continue to provide new opportunities in unveiling the complex processes underlying the basic building blocks of life. Molecular biology research has only just begun to study how proteins are spatially and temporally organized in larger functional units and how they behave under the influence of selective perturbations of the system by genetic and molecular interventions. Answering these questions will be critical to understanding diseases and our ability to design more effective drugs and therapeutic strategies. Since more and more research is being done in living cells, with time-lapse image data sets that are not only very large in size but

also highly variable and complex, research in this area is rapidly becoming dependent on automated techniques for image processing and analysis. Commercial software packages with modules for particle tracking and motion analysis are already available, but it is highly unlikely that a general-purpose algorithm, developed to provide a solution to many different tracking problems, is going to be the best fit for any particular tracking problem.

## Conclusions

In conclusion, we presented the combined theoretical and numerical investigation of the differential Zernike filter and its application in segmented telescope phasing and image processing. The results of our investigation are as follows:

Firstly, we studied a new differential Zernike filter based sensor (DZEUS) for phasing of a segmented mirror. We first presented in one dimension the analytical expression for the PtP values of output intensity modulations against the input phase jumps in DZEUS. We showed that the DZEUS gives rise to a better linear relation, compared to ZEUS, because the differential algorithm used in DZEUS removes the symmetrical (pollution) term that exists in the PtP expression of ZEUS. This makes DZEUS a better phase retrieval algorithm. In order to further improve the phasing accuracy, we put forward a multiple step correction approach which can further reduce the phase errors by iterations. Then, we extended the one dimension analysis to two dimensions and study the performance of DZEUS using numerically generated segmented mirrors. We showed that using the multiple step correction approach, the phasing error is reduced exponentially with respect to the number of iterations. We further compared the characteristics between DZEUS and ZEUS and concluded that the former performs better consistently. Finally, we showed that DZEUS is robust with respect to the atmospheric turbulence.

We further studied a phasing sensor system, which is realized by combining the differential Zernike filter and a feedback loop, and showed by analysis and simulation that this system provides an effective method to measure segment misalignment error in the presence of atmospheric turbulence. This phasing technique is different from DZEUS and ZEUS method. In ZEUS and DZEUS, only a single exposure from the Zernike filter output is used to retrieve the segmentation error based on its known wavefront-to-intensity converting relation for segmented mirror correction at each time. The atmospheric turbulence is partially averaged out by relatively long time exposure in ZEUS and DZEUS. While in the differential Zernike feedback system, the atmospheric turbulence is averaged out by multiple exposures.

The other part of this thesis focused on particle tracking in processing of living cell

images. We have shown by simulation that Zernike filtering can increase the tracking accuracy of a single particle tracking algorithm, in which, a local linking method is used, by suppressing the static object as well as the static noisy background, and makes the particle tracking algorithm work even with relatively weak signal (SNR as small as 3). Then, we conducted numerical experiments on three living cell image series, and show the effectiveness in a real application. In so far as the particles' density is low enough so that separately tracking of each particle is possible, this single particle tracking algorithm along with the Zernike filtering works effectively. However, in more complex situations with much higher densities and overlapping spots, the linking problem cannot be solved unambiguously locally. We employed an existing computational efficient and robust algorithm with couple of modification for two-dimensional multiple-particle tracking the particles in *Drosophila* oocyte image sequence, in which, the particle density is pretty high. In this tracking algorithm, a global linking technique based on graphic theory is employed. This algorithm works fairly well (with 89.5% linking successful rate, and no false link). We show, by comparison with manual tracking, this algorithm is limited as the tracking fails when particle intensity is too low.



## Appendix A

### Singular Value Decomposition

The measured intersegment phasing errors need to be transformed into piston values, so as to send to actuators to close the phasing loop. Now, the problem is how to transform the retrieved segment edge jumps into the corresponding piston value of the segment mirrors [77].

In our example, the primary mirror is consisted of 169 hexagonal segments, and there are 462 inter-segment edges. The piston value of each segment can be written as an one-dimensional matrix,  $u = \{P_1, P_2, \dots, P_{168}, P_{169}\}$ , while the inter-segment edge jump can be written as another one-dimensional matrix,  $v = \{J_1, J_2, \dots, J_{461}, J_{462}\}$ . There exist a  $462 \times 169$  two-dimensional matrix  $A$  that can make  $u$  and  $v$  linearly related:

$$Au = v. \quad (\text{A.1})$$

$B$  is the pseudoinverse matrix of matrix  $A$ , so we have:

$$u = Bv. \quad (\text{A.2})$$

The matrix  $B$  can be obtained by the following expression:

$$B = V\Sigma^+U^*, \quad (\text{A.3})$$

where  $V$  is a  $169 \times 169$  unitary matrix,  $\Sigma^+$  is the transpose of  $\Sigma$  with every nonzero entry replaced by its reciprocal, the matrix  $\Sigma$  is  $462 \times 169$  with nonnegative numbers on the diagonal and zero off the diagonal, and  $U^*$  denote the transpose of  $U$ , an  $462 \times 462$  unitary matrix.  $U$ ,  $\Sigma$ , and  $V$  are generated by singular value decomposition (SVD):

$$A = U\Sigma V^*, \quad (\text{A.4})$$

where  $V^*$  is the transpose of  $V$ .

## Appendix B

### Atmospheric Turbulence Simulation

The simulation of atmospherically distorted wavefront is important in the context of studies of light propagation and imaging through the atmosphere, as well as in the context of the correction of atmospherically distorted images where the distortion is treated as a disturbance, such as in an adaptive optics system. Research on the simulation of the atmosphere started with the general procedure described by McGlamery [42]: random phases are produced over the pupil, and the atmospheric correlation is introduced with proper filtering by the Kolmogorov spectrum [78].

A phase-screen realization over a grid of points is generated in the Fourier, or fast Fourier transform (FFT) method, by taking the FFT of filtered white noise. If  $\phi_{s,t}$  is the phase at the  $s$ -th row and  $t$ -th column of the grid of phase points, and if  $F$  is an  $N_y \times N_x$  array representing the phase power spectral density, then, according to the Fourier method [79],

$$\begin{pmatrix} \text{Re} \\ \text{Im} \end{pmatrix} \Phi_{s,t} = \begin{pmatrix} \text{Re} \\ \text{Im} \end{pmatrix} \left\{ \sum_{m=1}^{N_y} \sum_{n=1}^{N_x} \Phi_{m,n}^{1/2} [q_{m,n}^{(r)} + i q_{m,n}^{(i)}] \times \exp[i2\pi(m-1)(s-1)/N_y] \right. \\ \left. \times \exp[i2\pi(n-1)(t-1)/N_x] \right\}. \quad (\text{B.1})$$

In Eq. B.1,  $q_{m,n}^{(r)}$  and  $q_{m,n}^{(i)}$  are independent, Gaussian, random numbers with mean 0 and variance 1. That is, they obey the statistical properties

$$\begin{aligned} \langle q_{m,n}^{(r)} \rangle &= \langle q_{m,n}^{(i)} \rangle = 0, \\ \langle q_{m,n}^{(r)} q_{m',n'}^{(r)} \rangle &= \langle q_{m,n}^{(i)} q_{m',n'}^{(i)} \rangle = \delta_{m,m'} \delta_{n,n'}, \\ \langle q_{m,n}^{(r)} q_{m',n'}^{(i)} \rangle &= 0, \end{aligned} \quad (\text{B.2})$$

where

$$\delta_{m,m'} = \begin{cases} 1 & m = m' \\ 0 & \text{otherwise} \end{cases}. \quad (\text{B.3})$$

The operator (Real Imag) T in Eq. B.1 is needed because the two-dimensional discrete Fourier transform generates a “complex phase”. If we define  $\phi_{s,t} = \phi_{s,t}^{(r)} + i\phi_{s,t}^{(i)}$ , where

$\phi_{s,t}^{(r)}$  and  $\phi_{s,t}^{(i)}$  are real, then it can be shown that the phase screen with elements  $\phi_{s,t}^{(r)}$  is independent of that containing the elements  $\phi_{s,t}^{(i)}$ . The Fourier method is a very efficient algorithm for creating phase-screen realizations since the FFT can be employed. In addition, two independent phase screens are created per operation [ $\phi^{(r)}$  and  $\phi^{(i)}$ ], where  $\phi = \phi^{(r)} + i\phi^{(i)}$ .

It is straightforward to show that both  $\phi^{(r)}$  and  $\phi^{(i)}$ , as generated from Eq. B.1, satisfy the required property

$$\begin{aligned} \langle \phi_{s_1,t_1} \phi_{s_2,t_2} \rangle &= \frac{1}{2} \sum_{m=1}^{N_y} \sum_{n=1}^{N_x} \Phi_{m,n} (\exp\{i2\pi[(m-1)(s_2-s_1)/N_y + (n-1)(t_2-t_1)/N_x]\} \\ &+ \exp\{-i2\pi[(m-1)(s_2-s_1)/N_y + (n-1)(t_2-t_1)/N_x]\}). \end{aligned} \quad (\text{B.4})$$

For uncorrected phase errors, it follows from Kolmogorov theory that the power spectral density is [79]

$$\Phi(\kappa) d^2\kappa = 0.02288 r_0^{-5/3} \frac{d^2\kappa}{\kappa^{11/3}}, \quad (\text{B.5})$$

$r_0 \equiv$  Fried coherence diameter

For numerical analysis the discrete version of Eq. B.5 must be used. For a  $L_x \times L_y$  rectangular phase screen consisting of  $N_x \times N_y$  grid points (separation of  $L_x/N_x = L_y/N_y$  between adjacent phase grid points), it is straightforward to show that

$$\Phi_{m,n} = 0.02288 M_L (L_x/r_0)^{5/3} [M_L^2 \kappa_m^2 + \kappa_n^2]^{-11/6}, \quad (\text{B.6})$$

where

$$M_L \equiv L_x/L_y = N_x/N_y, \quad (\text{B.7})$$

$$\kappa_m \equiv \begin{cases} m-1 & m = (1, N_y/2) \\ -N_y-1+m & m = (N_y/2+1, N_y) \end{cases}, \quad (\text{B.8})$$

$$\kappa_n \equiv \begin{cases} n-1 & n = (1, N_x/2) \\ -N_x-1+n & n = (N_x/2+1, N_x) \end{cases}. \quad (\text{B.9})$$

In Eq. B.8 and B.9 we have assumed that  $N_y$  and  $N_x$  are even integers (typically they are powers of 2, since the FFT is used). The odd ordering of the discrete spatial

frequencies was, of course, needed so that they would coincide with the spatial frequencies of the term  $\exp[i2\pi(m-1)(s-1)/N_y]\exp[i2\pi(n-1)(t-1)/N_x]$  in Eq. B.1. (For example, for  $N=8$ , the spectrum would be ordered as (0, 1, 2, 3, -4, -3, -2, -1). Finally, we note that for  $m=n=1$  (dc term),  $\Phi_{0,0}$  is infinite, as calculated from Eq. B.6, B.8, and B.9. This pole is removed by choosing a normalization of 0 for the dc component of the discrete power spectral density,

$$\Phi_{0,0} = 0. \quad (\text{B.10})$$

A shortcoming of the Fourier method is that low spatial frequencies, which are the major contributor to uncorrected phase aberrations, are underrepresented in the phase screens. The reason for this is that for a phase screen of width  $L$ , the lowest spatial frequency in the discrete spectrum is  $1/L$ .

There are several methods for compensating for this underrepresentation due to the  $1/L$  cutoff. The most obvious scheme is to increase  $L$  so that  $L/D \gg 1$ , where  $D$  is the aperture diameter. This is not feasible for commonly used aperture diameters. For example, for  $L/D \gg 100$ ,  $D = 1\text{m}$ , and grid point sampling of 1 cm, the total phase screen would consist of an array of  $10,000 \times 10,000$  points. Processing arrays of this size presents computational difficulties. An alternative approach is to imbed the phase screen within a much larger screen whose grid point separation is equal to the full width of the screen being modelled. The larger screen can then be used to augment the low-frequency spectrum of the smaller screen. If the size of the larger coarse screen is  $N' \times N'$  grid points, then the spatial frequencies  $[1/N'L, 2/N'L, \dots, (N'-1)/N'L, 1/L]$  can be included in the modelling.

## Appendix C

### Single Particle Tracking Method

For methods based solely on centroid calculation, a search region is placed as tightly around the object of interest as possible (given how far the particle can move between frames). Before the centroid is calculated, a threshold operation is performed to remove as much background noise as possible without losing the edge of the object. Then the centre of mass is calculated and used to re-centre the search region for the next frame. The energy centre of the grayscale image of width  $w$  and height  $h$  is calculated as follows:

$$\bar{x} = \frac{\sum_{x=0}^w \sum_{y=0}^h xI(x, y)}{\sum_{x=0}^w \sum_{y=0}^h I(x, y)}, \quad (\text{C.1})$$

$$\bar{y} = \frac{\sum_{x=0}^w \sum_{y=0}^h yI(x, y)}{\sum_{x=0}^w \sum_{y=0}^h I(x, y)}, \quad (\text{C.2})$$

where  $x$  is the current x position,  $y$  is the current y position, and  $I(x, y)$  is the intensity of the image at point  $(x, y)$ . Various methods may be used to select an appropriate threshold. An appropriate threshold is one that minimizes the appearance of noise while retaining as much of the object (and its shape) as possible. There is no calibration with respect to particle shapes or other physical characteristics for this simple tracking method (calibrations by  $m_0$  and  $m_2$  are used in the multiple particle tracking).

Tracking a particle in an image sequence using a correlation method requires choosing a template image (kernel) to be searched for in subsequent images. The algorithm overlays and calculates a cross correlation with kernel for each point within the selected search region, creating a correlation image where higher brightness indicates higher correlation. A threshold is applied to the correlation image to remove the background and the centre of the correlation image is then found using a centroid interpolation method. For the cross-correlation, the kernel image is overlayed and compared to the image of interest and moved over the image of interest one pixel at a

time, with each location being scored for its similarity to the kernel image. This set of similarity scores makes a correlation matrix. The correlation matrix has the advantage of low noise, since more points are compared to many other points at the same time to create it. The normalized correlation coefficient is calculated as follows:

$$R(x, y) = \frac{\sum_{x'}^w \sum_{y'}^h (T(x', y') - \bar{T})(I(x + x', y + y') - \bar{I})}{\sqrt{\sum_{x'}^w \sum_{y'}^h (T(x', y') - \bar{T})^2 \sum_{x'}^w \sum_{y'}^h (I(x + x', y + y') - \bar{I})^2}} \quad (\text{C.3})$$

where  $R$  is the correlation image,  $w$  is the width of the template,  $h$  is the height of the template,  $T$  is the kernel image,  $I$  is the search area image,  $\bar{T}$  is the mean pixel value of the kernel,  $\bar{I}$  is the mean pixel value of the image within the current area  $(x', y')$  to  $(x'+w, y'+h)$ ,  $(x, y)$  and  $(x + x', y + y')$  are the pixel positions in the respective images (template and search area).

The correlation coefficient gives a result for each point between  $-1$  and  $1$ . Negative values indicate negative correlation and are excluded. Correlation methods on their own give a correlation matrix  $R(x, y)$ . By selecting the brightest point you can get an estimated accuracy to one pixel of the centre of the object, and sub-pixel accuracy can be realized by apply centroid interpolation to the correlation matrix.

## **Publications**

1. **Tao Huang**, Yingji He, Yang Xiao, Wenrui Xue, and Guosheng Zhou, “Effects of residual second-order dispersion to fourth-order dispersion in ultra high-speed optical time-division multiplexed transmission”, Chinese Opt. Lett. 12, 72 (2004).
2. Rongcao Yang, Wenrui Xue, **Tao Huang**, and Guosheng Zhou, “Research on the effects of air hole shape on the properties of microstructured optical fibers”, Opt. Eng. 43, 2701 (2004).
3. **T. Huang**, W. Lu, S. Zhang, A. H. Greenaway, “Zernike phase sensor for phasing of segment telescopes,” Appl. Phys. B 86, 139-145 (2007).
4. **T. Huang**, W. Lu, A. H. Greenaway, “Differential Zernike phase sensor for phasing of the segmented mirror: theory and comparison with ZEUS”, manuscript prepared.

## References

1. J. E. Nelson, T. S. Mast, and S. M. Faber, "The design of the Keck Observatory and telescope," Keck Observatory Report **90**, 5-1 to 5-44 (1985).
2. N. Yaitskova and K. Dohlen, "Tip-tilt error for extremely large segmented telescopes: detailed theoretical point-spread function analysis and numerical simulation results," JOSA A. **19**, 1274-1285 (2002).
3. N. Yaitskova and K. Dohlen, "Simulation of imaging performance for extremely large segmented telescopes," in *Optical Design, Materials, Fabrication, and Maintenance*, Proc. SPIE **4003**, 279-290 (2000).
4. G. Zeider and E. Montgomery, "Diffraction effect with segmented aperture," in *Space Telescopes and Instruments V*, Proc. SPIE **3356**, 799-809 (1998).
5. T. S. Mast, J. E. Nelson, "Segmented mirror control system hardware for CELT," in *Optical Design, Materials, Fabrication, and Maintenance*, Proc. SPIE **4003**, 226-240 (2000).
6. G. A. Chanan, "Design of the Keck Observatory alignment camera," in *Precision Instrument Design*, SPIE **1036**, pp. 59-70 (1988).
7. G. A. Chanan, J. Nelson, T. Mast, P. Wizinowich, and B. Schaefer, "The W. M. Keck Telescope phasing camera system," in *Instrumentation in Astronomy VIII*, SPIE **2198**, 1139-1150 (1994).
8. G. Chanan, M. Troy, F. Dekens, S. Michaels, J. Nelson, T. Mast, and D. Kirkman, "Phasing the Mirror Segments of the Keck Telescopes: The Broadband Phasing Algorithm," Appl. Opt. **37**, 140-155 (1998).
9. G. A. Chanan, M. Troy, and E. Sirko, "Phase discontinuity sensing: A method for phasing segmented mirrors in the infrared," Applied Optics **38**, 704-713 (1999).
10. S. Cuevas, V. G. Orlov, F. Garfias, V. Voitsekhovich, and L. Sanchez, "Curvature equation for a segmented telescope," in *Optical Design, Materials, Fabrication, and Maintenance*, Proc. SPIE **4003**, 291-302 (2000).
11. L. Montoya, N. Yaitskova, P. Dierickx, and K. Dohlen, "Mach-Zehnder wave front sensor for phasing of segmented telescopes," in *Future Giant Telescopes*, Proc. SPIE **4840**, 564-573 (2002).



12. N. Yaitskova, K. Dohlen, P. Dierichx, L. Montoya, "Mach-Zehnder interferometer for piston and tip-tilt sensing in segmented telescopes: theory and analytical treatment," *J. Opt. Soc. Am. A* **22**, 1093 (2005).
13. S. Esposito, E. Pinna, A. Tozzi, P. Stefanini, and N. Devaney, "Co-phasing of segmented mirrors using pyramid sensor," in *Astronomical Adaptive Optics Systems and Applications*, Proc. SPIE **5169**, 214–225 (2003).
14. S. Esposito, E. Pinna, A. Puglisi, A. Tozzi, and P. Stefanini, "Pyramid sensor for segmented mirror alignment," *Opt. Lett.* **30**, 2572-2574 (2005).
15. K. Dohlen, M. Langlois, P. Lanzoni, S. Mazzanti, A. Vigan, L. Montoya, E. Hernandez, M. Reyes, I. Surdej, N. Yaitskova, "ZEUS: A Cophasing Sensor Based on the Zernike Phase Contrast Method," Proc. SPIE **6267**, 102 (2006).
16. J. Schwiegerling, and D. R. Neal, "Historical Development of the Shack-Hartmann Wavefront Sensor," <http://www.wavefrontsciences.com/Historical%20Development.pdf>.
17. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," *Appl. Opt.* **27**, 1223-1225 (1988).
18. F. Roddier, "Wavefront sensing and the irradiance transport equation," *Appl. Opt.* **29**, 1402-1403 (1990).
19. N. Roddier, "Algorithms for wavefront reconstruction out of curvature sensing data," in *Active and Adaptive Optical Systems*, Proc. SPIE **1542**, 120-129 (1991).
20. G. Chanan, M. Troy, and C. Ohara, "Phasing the Primary Mirror Segments of the Keck Telescopes: A Comparison of Different Techniques," Proc. SPIE 2000.
21. R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," *Journal of Modern Optics* **43**, 289-293 (1996).
22. S. Esposito, N. Devaney, "Segmented telescopes co-phasing using Pyramid Sensor," Proc. Beyond Conventional Adaptive Optics, ESO Conference and Workshop Proceedings **58**, 161 (2002).
23. R. G. Lyon, "DCATT wavefront sensing and optical control study," Tech. Rep. WFSC-0001, NASA/GSFC (1999).
24. M. G. Lofdahl, R. L. Kendrick, A. Harwit, K. E. Mitchell, A. L. Duncan, J. H. Seldin, R. G. Paxman, and D. S. Acton, "A phase diversity experiment to measure

- piston misalignment on the segmented primary mirror of the Keck II telescope,” in Space Telescopes and Instrument V, Proc. SPIE **3356**, 1190-1201 (1998).
25. D. C. Ghiglia and M. D. Pritt, Two-Dimensional Phase Unwrapping, Wiley, New York, 1998.
  26. M. G. Lofdahl, and H. Eriksson, “Resolving piston ambiguities when phasing a segmented mirror,” in Optical, and IR Space Telescopes and Instruments VI, Proc. SPIE **4013**, 55 (2000).
  27. N. Yaitskova, F. Gonte, F. Derie, L. Noethe, I. Surdej, R. Karban, K. Dohlen, M. Langlois, S. Esposito, E. Pinna, M. Reyes, L. Montoya, and D. Terrett, “The active phasing experiment part I: concept and objectives,” in Ground Based and Airborne Telescopes, Proc. SPIE **6267**, 84 (2006).
  28. F. Gonte et al “The active phasing experiment. Part II: design and development,” in Ground Based and Airborne Telescopes, Proc SPIE **6267**, 85 (2006).
  29. M. A. Vorontsov, E. W. Justh, and L. A. Beresnev, “Adaptive optics with advanced phase-contrast techniques. I. High-resolution wave-front sensing,” J. Opt. Soc. Am. A **18**, 1289 (2001).
  30. E. W. Justh, M. A. Vorontsov, G. W. Carhart, L. A. Beresnev, and P. S. Krishmarasad, “Adaptive optics with advanced phase-contrast techniques. II. High-resolution wave-front control,” J. Opt. Soc. Am. A **18**, 1300 (2001).
  31. F. Zernike, “How I discovered phase contrast,” Science **121**, 345 (1955).
  32. J. W. Goodman, Introduction to Fourier Optics, McGrawHill, New York 1996.
  33. S. A. Akhmanov and S. Yu. Nikitin, Physical Optics, Clarendon, Oxford, UK, 1997.
  34. V. P. Sivokon and M. A. Vorontsov, “High-resolution adaptive phase distortion suppression based solely on intensity information,” J. Opt. Soc. Am. A **15**, 234–247 (1998).
  35. M. A. Vorontsov, “High-resolution adaptive phase distortion compensation using a diffractive-feedback system: experimental results,” J. Opt. Soc. Am. A **16**, 2567–2573 (1999).

36. R. A. Muller and A. Buffington, "Real-time correction of atmospherically degraded telescope images through image sharpening," *J. Opt. Soc. Am.* **64**, 1200–1210 (1974).
37. T. R. O'Meara, "The multi-dither principle in adaptive optics," *J. Opt. Soc. Am.* **67**, 306–315 (1977).
38. M. A. Vorontsov and V. I. Shmalhauzen, *Principles of Adaptive Optics* (Nauka, Moscow, 1985).
39. B. M. ter Haar Romey, ed., *Geometry-Driven Diffusion in Computer Vision* (Kluwer-Academic, Dordrecht, The Netherlands, 1994).
40. IEEE Trans. Image Process. Special issue on partial differential equations and geometry-driven diffusion in image processing and analysis, *IEEE Trans. Image Process.* **7**, (1998).
41. M. A. Vorontsov, "Parallel image processing based on an evolution equation with anisotropic gain: integrated optoelectronic architectures," *J. Opt. Soc. Am. A* **16**, 1623–1637 (1999).
42. B. L. McGlamery, "Computer simulation studies of compensation of turbulence degraded images," *Proc. SPIE* **74**, 225 (1976).
43. D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *J. Opt. Soc. Am.* **56**, 1372–1379 (1966).
44. N. Roddier, "Atmospheric wavefront simulation using Zernike polynomials," *Opt. Eng.* **29**, 1174–1180 (1990).
45. N. Roddier, "Atmospheric wavefront simulation using Zernike polynomials," *Opt. Eng.* **29**, 1174 (1990).
46. Chester S. Cardner, Byron M. Welsh, and Laird A. Thompson, "Design and performance analysis of adaptive optical telescopes using lasing guide stars," *Proc. IEEE* **78**, 1721 (1990).
47. Wilson R.W., "SLODAR: Measuring optical turbulence altitude with a Shack--Hartmann wave-front sensor", *MNRAS* **337**, 103 (2002).
48. G. Chanan, C. Ohara, and M. Troy, "Phasing the mirror segments of the Keck telescopes II: the narrow phasing algorithm," *Appl. Opt.* **39**, 4706-4714 (2000).

49. A. Schumacher, N. Devaney, and L. Montoya, "Phasing segmented mirrors: a modification of the Keck narrow-band technique and its application to extremely large telescopes," *Appl. Opt.* **41** 1297-1307 (2002).
50. Fang Shi, Gary Chanan, Catherine Ohara, Mitchell Troy, David C. Redding, "Experimental Verification of Dispersed Fringe Sensing as a Segment Phasing Technique using the Keck Telescope" *Appl. Opt.* **43**, 4474 (2004).
51. N. Yaitskova, "Performance analysis of Mach-Zehnder interferometer for detection of wave-front discontinuities," *Proc. SPIE* **5169**, 62 (2003).
52. J. N. Aubrun, K. R. Lorell, T. W. Havas, and W. C. Henninger, "Performance analysis of the segmented alignment control system for the ten-meter telescope," *Automatica* **24**, 437-454 (1988).
53. R. Cohen, T. Mast, and J. Nelson, "Performance of the W. M. Keck Telescope active mirror control system," *SPIE* **2199**, 105-116 (1994).
54. E. Meijering, I. Smal, and G. Danuser, "Tracking in Molecular Bioimaging", *IEEE SIGNAL PROCESSING MAGAZINE* **46**, 46-53 (2006).
55. R. Eils and C. Athale, "Computational imaging in cell biology," *The Journal of Cell Biology*, **161**, 477-481 (2003).
56. D. Gerlich, J. Mattes, and R. Eils, "Quantitative motion analysis and visualization of cellular structures," *Methods*, **29**, 3-13 (2003).
57. D. J. Stephens and V. J. Allan, "Light microscopy techniques for live cell imaging," *Science* **300**, 82-86 (2003).
58. J. Lippincott-Schwartz and G. H. Patterson, "Development and use of fluorescent protein markers in living cells," *Science* **300**, no. 5616, 87-91 (2003).
59. M. Gu, *Advanced Optical Imaging Theory*. Berlin: Springer-Verlag, 2000.
60. M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophys. J.* **81**, 2378-2388 (2001).
61. D. S. Martin, M. B. Forstner, and J. A. Käs, "Apparent subdiffusion inherent to single particle tracking," *Biophys. J.* **83**, 2109-2117 (2002).

62. H. Bornfleth, P. Edelmann, D. Zink, T. Cremer, and C. Cremer, "Quantitative motion analysis of subchromosomal foci in living cells using four-dimensional microscopy," *Biophys. J.* **77**, 2871–2886 (1999).
63. W. Tvaruskó, M. Bentele, T. Misteli, R. Rudolf, C. Kaether, D. L. Spector, H. H. Gerdes, and R. Eils, "Time-resolved analysis and visualization of dynamic processes in living cells," *Proc. Nat. Acad. Sciences Amer.* **96**, 7950–7955 (1999).
64. D. Uttenweiler, C. Weber, B. Jähne, R. H. A. Fink, and H. Scharr, "Spatiotemporal anisotropic diffusion filtering to improve signal-to-noise ratios and object restoration in fluorescence microscopic image sequences," *J. Biomed. Optics* **8**, 40–47 (2003).
65. C.J. Veenman, M.J.T. Reinders, and E. Backer, "Resolving motion correspondence for densely moving points," *IEEE Trans. Pattern Anal. Machine Intell.* **23**, 54–72 (2001).
66. B. C. Carter, G. T. Shubeita, and S. P. Gross, "Tracking single particles: a user-friendly quantitative evaluation," *Phys. Biol.* **2**, 60–72 (2005).
67. Ryan, T. A., P. J. Millard, and W. W. Webb, "Imaging  $[Ca^{2+}]$  dynamics during signal transduction," *Cell Calcium*. **11**, 145–155 (1990).
68. J. H. Simonetti, "Measuring the signal to noise ratio S/N of the CCD image of a star or nebula," <http://www.phys.vt.edu/~jhs/phys3154/snr20040108.pdf> (2004).
69. I. F. Sbalzarini, P. Koumoutsakos, "Feature point tracking and trajectory analysis for video imaging in cell biology," *Journal of Structural Biology* **151** 182–195 (2005).
70. A. Reilein, S. Yamada, and W. J. Nelson, "Self-organization of an acentrosomal microtubule network at the basal cortex of polarized epithelial cells", *JCB*. **171**, 845 (2005).
71. M. Piehl and L. Cassimeris, "Organization and Dynamics of Growing Microtubule Plus Ends during Early Mitosis", *Mol. Biol. Cell*, **14**, 916 (2003).
72. S. B. Dalziel, "Rayleigh–Taylor instability: experiments wit image analysis," *Dynam. Atmos. Oceans* **20**, 127–153 (1993).

73. S. B. Dalziel, "Decay of rotating turbulence: some particle tracking experiments," In: Nieuwstadt, F.T.M. (Ed.), *Flow Visualization And Image Analysis*. Kluwer, Dordrecht, 27–54 (1993).
74. F. L. Hichcock, "The distribution of a product from several sources to numerous localities," *J. Math. Phys.* **20**, 224 (1941).
75. D. Chetverikov, J. Verestoy , "Feature point tracking for incomplete trajectories," *Computing* **62**, 321–338 (1999).
76. J. C. Crocker, D.G. Grier, "Methods of digital video microscopy for colloidal studies," *J. Coll. Interface Sci.* **179**, 298–310 (1996).
77. N. Yaitskova, L. Montoya-Martinez, K. Dohlen, and P. Dierickx, "A Mach-Zehnder phasing sensor for extremely large segmented telescopes: laboratory results and close loop algorithm," in *Ground-Based Telescopes*, Proc. SPIE **5489**, 1139–1151 (2004).
78. Jeffrey B. Shellan, "Statistical properties of the Strehl ratio as a function of pupil diameter and level of adaptive optics correction following atmospheric propagation," *J. Opt. Soc. Am. A* **21**, 1445-1451 (2004).
79. R. G. Lane, A. Glindemann, and J. C. Dainty, "Simulation of a Kolmogorov phase screen," *Waves Random Media* **2**, 209–224 (1992).
80. A. K. Jain, "Fundamentals of Image Processing," Prentice Hall, Englewood Cliffs, NJ, 1986.
81. R. C. Gonzalez, and P. Wintz, "Digital Image Processing, 2nd Edition," Addison–Wesley, Reading, MA, 1987.
82. W. K. Pratt, "Digital Image Processing," Wiley, New York, 1991.